

Using Coalgebras and the Giry Monad for Interpreting Game Logics — A Tutorial

Ernst-Erich Doberkat*
 MATH ++ SOFTWARE, Bochum, Germany
 math@doberkat.de

January 3, 2017

Contents

1 Objectives	2
2 Coalgebras	2
3 The $\\$-Functor	4
4 The Giry Monad	9
5 Playing Around with Morphisms	15
6 Bisimulations	18
7 Expressivity of Kripke Models	23
8 Stochastic Effectivity Functions	28
9 What To Do Next	38
A The π-λ-Theorem	39
B The Souslin Operation	39

Abstract

The stochastic interpretation of Parikh’s game logic should not follow the usual pattern of Kripke models, which in turn are based on the Kleisli morphisms for the Giry monad, rather, a specific and more general approach to probabilistic nondeterminism is required. We outline this approach together with its probabilistic and measure theoretic basis, introducing in a leisurely pace the Giry monad and their Kleisli morphisms together with important techniques for manipulating them. Proof establishing specific techniques are given, and pointers to the extant literature are provided.

After working through this tutorial, the reader should find it easier to follow the original literature in this and related areas, and it should be possible for her or him to appreciate measure theoretic arguments for original work in the areas of Markov transition systems, and stochastic effectivity functions.

*Some of the results were obtained while the author held the Chair for Software Technology at Technische Universität Dortmund; they were funded in part by Deutsche Forschungsgemeinschaft, *Koalgebraische Eigenschaften stochastischer Relationen*, grant DO 263/12-1.

AMS subject classification: 03B45, 18C15, 18C20

Computing Reviews Classification: F.4.1, I.2.3, I.2.4, G.3

1 Objectives

A minimal categorial framework is introduced in order to formulate coalgebras and monads (which come here in their disguise as Kleisli tripels). We specialize then to the category of measurable spaces, discussing here in particular the Giry monad, with occasional side glances to the upper closed functor. This is complemented by a discussion of morphisms for stochastic coalgebras (which will also be used for the interpretation of various modal logics), giving among others congruences, which will be put to use when discussing the expressivity of Kripke models. But before doing so, we have a fairly general look at bisimulations for various transition models, pointing at difficulties arising for stochastic coalgebras. We expand then our scenario by introducing stochastic effectivity functions, which we will briefly investigate, and which will be used for an interpretation of game logics.

Not all topics can be treated in depth due to limitations in space, but proofs are provided here and there, mostly for illustrating some techniques. Two appendices are provided, one discusses a technical device (the Souslin operation), the other one gives the important π - λ -Theorem from Boolean σ -algebras.

The classic reference to coalgebras is the paper by Rutten [21], the survey paper by Venema [26] focusses on representation issues, see also [15]. The present discussion is based on [9, 8, 10]. References to extant pieces of literature are given when needed.

2 Coalgebras

Fix a category \mathbf{C} with an endofunctor \mathbf{F} (I assume that the reader knows what a category is, and what a functor does).


Definition 2.1 An \mathbf{F} -coalgebra (\mathbf{a}, f) over \mathbf{C} is an object \mathbf{a} of \mathbf{C} together with a morphism $f : \mathbf{a} \rightarrow \mathbf{F}\mathbf{a}$.

Coalgebra

Example 2.2 Let \mathbf{C} be the category of sets with maps as morphisms, the functor is the power set functor $\mathbf{2}^-$. (A, f) is an $\mathbf{2}^-$ -coalgebra iff $f : A \rightarrow \mathbf{2}^A$ is a map. This is in 1-1-correspondence with binary relations:

Category of sets

$$\langle x, x' \rangle \in R \text{ iff } x' \in f(x).$$

Through this, *transition systems* are studied. 

Example 2.3 Let X resp. Y be the inputs and the outputs of an automaton with outputs. Define $\mathbf{F} := (- \times Y)^X$ over the category of sets. The (A, f) is an \mathbf{F} -coalgebra iff it is an automaton with states A : Since $f : A \rightarrow (A \times Y)^X$, we have $f(\mathbf{a})(x) \in A \times Y$, say, $f(\mathbf{a})(x) = \langle \mathbf{a}', y \rangle$, hence \mathbf{a}' is the new state of the automaton, y its output upon input x in state \mathbf{a} .

Automata

Conversely, let (X, Y, A, δ) be an automaton with output, i.e., $\delta : A \times X \rightarrow A \times Y$ is a map. Currying gives a map $f : A \rightarrow (A \times Y)^X$ through $f(a)(x) := \delta(a, x)$. This means that f is an \mathbf{F} -coalgebra. \uparrow

Example 2.4 Put $\mathbf{F}A := \{*\} \cup A \times A$ with $*$ a new symbol. (A, f) is an \mathbf{F} -coalgebra iff f corresponds to a binary tree over A . Put $f(a) := *$ iff a is a leaf, and $f(a) = \langle a_1, a_2 \rangle$ iff a_1 and a_2 are offsprings of a . The tree may be infinite, though. \uparrow

Trees

Fix in what follows both \mathbf{C} and \mathbf{F} .

Definition 2.5 Let (a_1, f_1) and (a_2, f_2) be \mathbf{F} -coalgebras. A \mathbf{C} -morphism $\varphi : a_1 \rightarrow a_2$ is a coalgebra morphism $(a_1, f_1) \rightarrow (a_2, f_2)$ iff $f_2 \circ \varphi = \mathbf{F}\varphi \circ f_1$. This means that the diagram

Coalgebra morphism

$$\begin{array}{ccc} a_1 & \xrightarrow{\varphi} & a_2 \\ f_1 \downarrow & & \downarrow f_2 \\ \mathbf{F}a_1 & \xrightarrow{\mathbf{F}\varphi} & \mathbf{F}a_2 \end{array}$$

commutes.

Proposition 2.6 \mathbf{F} -coalgebras form a category with coalgebra morphisms as morphisms; composition is inherited from \mathbf{C} . \dashv

Example 2.7 Consider the coalgebras corresponding to transition systems from Example 2.2. Then $\varphi : (A_1, f_1) \rightarrow (A_2, f_2)$ is a coalgebra morphism iff these conditions are satisfied:

1. $a'_1 \in f_1(a_1)$ implies $\varphi(a'_1) \in f_2(a_1)$.
2. If $a'_2 \in f_2(\varphi(a_1))$, then there exists $a'_1 \in f_1(a_1)$ with $\varphi(a'_1) = a'_2$.

In fact, assume that φ is a coalgebra morphism. Then we note that $\mathbb{Z}^\varphi(W) = \varphi[W]$, and that we have this commuting diagram

$$\begin{array}{ccc} A_1 & \xrightarrow{\varphi} & A_2 \\ f_1 \downarrow & & \downarrow f_2 \\ \mathbb{Z}^{A_1} & \xrightarrow{\mathbb{Z}^\varphi} & \mathbb{Z}^{A_2} \end{array}$$

Assume that $a'_1 \in f_1(a_1)$, then

$$\varphi(a'_1) \in \varphi[f_1(a_1)] = \mathbb{Z}^\varphi(f_1(a_1)) = f_2(\varphi(a_1)).$$

This gives us the first condition. On the other hand, let $a'_2 \in f_2(\varphi(a_1)) = \mathbb{Z}^\varphi(f_1(a_1))$, thus $a'_2 \in \varphi[f_1(a_1)]$. But this implies that we can find $a'_1 \in f_1(a_1)$ with $\varphi(a'_1) = a'_2$. This provides us with the second condition. The converse direction offers itself as an exercise.

Bounded morphisms

These morphisms are called *bounded morphisms* in the theory of transition systems. In fact, reformulate $a' \in f_1(a)$ as $a \rightarrow_1 a'$, similarly $b' \in f_2(b)$ as $b \rightarrow_2 b'$. In this notation, the conditions above read

1. $a_1 \rightarrow_1 a'_1$ implies $\varphi(a_1) \rightarrow_2 \varphi(a'_1)$.
2. If $\varphi(a_1) \rightarrow_2 a'_2$, then there exists a'_1 with $a_1 \rightarrow_1 a'_1$ such that $\varphi(a'_1) = a'_2$.

This indicates an interesting connection between coalgebras and transition systems. 

Upper
closed

Example 2.8 Take the functor \mathbb{E} of all upper closed subsets of $\mathbf{2}^+$. We did not yet define how \mathbb{E} acts on maps, hence we have to transform $f : A \rightarrow B$ to $\mathbb{E}(f) : \mathbb{E}(A) \rightarrow \mathbb{E}(B)$. Let's see how to do this.

Let $\mathcal{G} \in \mathbb{E}(A)$, then $\mathcal{G} \subseteq \mathbf{2}^A$ is upper closed. Thus the set $\{H \subseteq B \mid f^{-1}[H] \in \mathcal{G}\} \subseteq \mathbf{2}^B$ is also upper closed (if $H_1 \subseteq H_2$ and $f^{-1}[H_1] \in \mathcal{G}$, we note that $f^{-1}[H_1] \subseteq f^{-1}[H_2]$, and since \mathcal{G} is upper closed, this implies that $f^{-1}[H_2] \in \mathcal{G}$). With this in mind, we put

$$\mathbb{E}(f)(\mathcal{G}) := \{H \subseteq B \mid f^{-1}[H] \in \mathcal{G}\}.$$

It can be shown that $\mathbb{E}(g \circ f) = \mathbb{E}(g) \circ \mathbb{E}(f)$ (see [9, 2.3.14]).

What do morphisms for \mathbb{E} look like? Let's try:


$$\begin{array}{ccc} A_1 & \xrightarrow{\varphi} & A_2 \\ f_1 \downarrow & & \downarrow f_2 \\ \mathbb{E}(A_1) & \xrightarrow{\mathbb{E}(\varphi)} & \mathbb{E}(A_2) \end{array}$$

Let $\mathcal{H} = f_2(\varphi(a_1)) = (\mathbb{E}(\varphi) \circ f_1)(a_1)$, thus

$$H \in \mathcal{H} \text{ iff } H \in \mathbb{E}(\varphi)(f_1(a_1)) \text{ iff } \varphi^{-1}[H] \in f_1(a_1).$$

Consequently we have

$$f_2(\varphi(a_1)) = \{H \subseteq B \mid \varphi^{-1}[H] \in f_1(a_1)\}$$

as a qualifying condition for $\varphi : A_1 \rightarrow A_2$ to become a \mathbb{E} -morphism. 

3 The \mathbb{S} -Functor

Before entering the a discussion on the probability functor, we need to know a little bit more about measurable spaces. Recall that a measurable space (X, \mathcal{A}) is a set X together with a Boolean σ -algebra \mathcal{A} on X , the σ indicating here that the Boolean algebra is closed under countable unions (and, by implication, under countable intersections).

Definition 3.1 Let X be a set, $\mathcal{A} \subseteq \mathbf{2}^X$ be a family of subsets of X . Then

$\sigma(\mathcal{A})$

$$\sigma(\mathcal{A}) := \bigcap \{B \mid \mathcal{A} \subseteq B, B \text{ is a } \sigma\text{-algebra}\}$$

is the smallest σ -algebra on X which contains \mathcal{A} . \mathcal{A} is called a generator of $\sigma(\mathcal{A})$.

It is clear that $\sigma(\mathcal{A})$ is always a σ -algebra (check the properties). Also, $\mathcal{A} \subseteq \mathbf{2}^X$, the latter one being a σ -algebra. Note that $\sigma : \mathbf{2}^X \rightarrow \mathbf{2}^X$ is a closure operator, thus we have

- if $\mathcal{A} \subseteq \mathcal{B}$, then $\sigma(\mathcal{A}) \subseteq \sigma(\mathcal{B})$,
- $\mathcal{A} \subseteq \sigma(\mathcal{A})$,
- $\sigma(\sigma(\mathcal{A})) = \sigma(\mathcal{A})$. In particular, $\sigma(\mathcal{A}) = \mathcal{A}$, whenever \mathcal{A} is a σ -algebra.

Example 3.2 The *Borel sets* $\mathcal{B}(\mathbb{R})$ are defined as the smallest σ -algebra on \mathbb{R} which contains the open (or the closed) sets. We claim that

Borel sets

$$\mathcal{B}(\mathbb{R}) = \sigma(\{[r, +\infty[\mid r \in \mathbb{R}\}) =: \mathcal{Q}.$$

In fact

- $\mathcal{Q} \subseteq \mathcal{B}(\mathbb{R})$, since

$$[r, +\infty[= \bigcap_{n \in \mathbb{N}} [r - \frac{1}{n}, \infty[,$$

the latter sets are in $\mathcal{B}(\mathbb{R})$, since they are open.

- $[r, s[\in \mathcal{Q}$ for $r < s$, since $[r, s[= [r, \infty[\setminus [s, \infty[$.
- $]r, s[\in \mathcal{Q}$, since $]r, s[= \bigcup_{n \in \mathbb{N}} [r + 1/n, s[$.
- All open sets in \mathbb{R} are in \mathcal{Q} , because each open set can be written as the union of countably many open intervals (and the open intervals are in \mathcal{Q}).
- All closed sets are in \mathcal{Q} as well, since \mathcal{Q} is closed under complementation. Hence $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{Q}$.



The reals \mathbb{R} are **always** assumed to have the Borel σ -algebra. In what follows, we will usually write down measurable spaces without their σ -algebras, unless we have to.

Definition 3.3 Let (X, \mathcal{A}) and (Y, \mathcal{B}) be measurable spaces. A map $f : X \rightarrow Y$ is called \mathcal{A} - \mathcal{B} -measurable iff $f^{-1}[B] \in \mathcal{A}$ for all $B \in \mathcal{B}$.

Measurability

Note the similarity to continuity (*inverse images of open sets are open* is the general definition), and to uniform continuity (resp. *inverse images of neighborhoods are neighborhoods*). Note also that a measurable map is not necessarily a Boolean homomorphism of the Boolean algebras \mathcal{A} and \mathcal{B} (constant maps are always measurable, but are rarely homomorphisms of Boolean algebras).

We should convince ourselves that we did indeed create a category.

Proposition 3.4 Measurable spaces with measurable maps form a category.

Proof We need only to show: if $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ are measurable, so is $g \circ f : X \rightarrow Z$. Let \mathcal{A} , \mathcal{B} , and \mathcal{C} be the corresponding σ -algebras, then we obtain from the measurability of g that $g^{-1}[C] \in \mathcal{B}$ for $C \in \mathcal{C}$, thus $f^{-1}[g^{-1}[C]] \in \mathcal{A}$, since f is also measurable. But $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$, so the assertion follows. \dashv

This is a helpful criterion for measurability, since it permits testing only on a generator, which is usually more readily available than the whole σ -algebra.

Lemma 3.5 *Let (X, \mathcal{A}) and (Y, \mathcal{B}) be measurable spaces, $f : X \rightarrow Y$ a map, and assume that \mathcal{B} is generated by \mathcal{B}_0 . Then f is \mathcal{A} - \mathcal{B} -measurable iff $f^{-1}[B] \in \mathcal{A}$ for all $B \in \mathcal{B}_0$.*

Proof 1. The condition is clearly necessary for measurability, since $\mathcal{B}_0 \subseteq \mathcal{B}$.

2. The criterion is also sufficient. We show this through the *principle of good sets* (see [9, Remark after Theorem 1.6.30]). It works like this. We want to show that $f^{-1}[B] \in \mathcal{A}$ holds for all $B \in \mathcal{B}$. Consider the set \mathcal{G} of all “good sets”,

$$\mathcal{G} := \{B \subseteq Y \mid f^{-1}[B] \in \mathcal{A}\}.$$

Then

1. \mathcal{G} is a σ -algebra. This is so because f^{-1} is compatible with all the Boolean operations, e.g., $f^{-1}[\bigcup_{i \in I} B_i] = \bigcup_{i \in I} f^{-1}[B_i]$.
2. $\mathcal{B}_0 \subseteq \mathcal{G}$ by assumption.

Thus

$$\sigma(\mathcal{B}_0) \subseteq \sigma(\mathcal{G}) = \mathcal{G}.$$

Hence $\mathcal{B} = \sigma(\mathcal{B}_0) \subseteq \mathcal{G}$, but this means that $f^{-1}[B] \in \mathcal{A}$ for all $B \in \mathcal{B}$. Hence f is in fact \mathcal{A} - \mathcal{B} -measurable. \dashv

This is an easy consequence from Lemma 3.5 together with Example 3.2:

Lemma 3.6 *A map $f : X \rightarrow \mathbb{R}$ is measurable iff $\{x \in X \mid f(x) \geq r\}$ is a measurable subset of X for all $r \in \mathbb{R}$. \dashv*

Note that we can replace the sets $\{f \geq r\}$ by $\{f \leq r\}$, by $\{f > r\}$ or by $\{f < r\}$, since there are easy ways to compute one through the other using only countable operations (such as $\{f > r\} = \bigcup_{n \in \mathbb{N}} \{f \geq r + 1/n\}$), and so on.

Example 3.7 Let χ_A be the *indicator function* of $A \subseteq X$, thus

$$\chi_A(x) := \text{if } x \in A \text{ then } 1 \text{ else } 0 \text{ fi.}$$

Then χ_A is a measurable function iff A is a measurable set. This becomes evident from

$$\{x \in X \mid \chi_A(x) \geq r\} = \begin{cases} \emptyset, & r > 1, \\ A, & 0 < r \leq 1, \\ X, & r \leq 0 \end{cases}$$

A measurable space induces a measurable structure on the space

$$\$(X) := \{\mu \mid \mu \text{ is a probability on (the } \sigma\text{-algebra of) } X\}$$

in the following way. Define first

$$\beta(A, q) := \{\mu \in \$(X) \mid \mu(A) \geq q\}$$

as the set of measures the value of which at event A is not smaller than q .

Definition 3.8 *Given a measurable space X , its \ast - σ -algebra is the smallest σ -algebra on $\$(X)$ which contains the sets $\{\beta(A, q) \mid A \subseteq X \text{ measurable, } q \in \mathbb{R}\}$.*

\ast - σ -algebra

Thus the \ast - σ -algebra is the smallest σ -algebra on $\$(X)$ which produces measurable maps from the evaluation at events.

Example 3.9 Define for the measurable space X the embedding $\eta_X : X \rightarrow \$(X)$ upon setting $\eta_X(x)(A) := \chi_A(x)$. Then η_X is a measurable map. In fact, by Lemma 3.5 and the definition of the \ast - σ -algebra we have to show that the set

$$\eta_X^{-1}[\beta(A, q)] = \{x \in X \mid \eta_X(x) \in \beta(A, q)\}$$

is measurable for each measurable set $A \subseteq X$ and each $q \in \mathbb{R}$. But we have

$$\eta_X(x) \in \beta(A, q) \text{ iff } \chi_A(x) \geq q,$$

so that the set in question is measurable by Example 3.7. η is usually called the *Dirac kernel*, $\eta_X(x)$ the *Dirac measure* on $x \in X$, which is usually denoted by δ_x , when the measurable space X is understood.

Reformulating, we see that $\eta_X : X \rightarrow \$X$ is a morphism in the category of measurable spaces. \mathbb{M}

We define now $\$$ for a measurable map and show that this yields a measurable map again. This is the basis for

1. showing that $\$$ is an endofunctor on the category of measurable spaces,
2. establishing the properties of the Giry monad.

Allora:

Definition 3.10 *Let X and Y be measurable spaces, $f : X \rightarrow Y$ be a measurable map. Define for $\mu \in \$(X)$ and for $B \subseteq Y$ measurable*

$$\$(f)(\mu)(B) := \mu(f^{-1}[B]).$$

This is called the image measure for μ under f .

Image
measure

The first step towards showing that $\$$ is an endofunctor consists in showing that $\$$ transforms measurable maps into measurable maps again (albeit on another space).

Lemma 3.11 *Given X, Y and f as above, $\$(f) : \$ (X) \rightarrow \$ (Y)$ is measurable with respect to the * - σ -algebras.*

Proof 0. We have to establish first that $\$(f)(\mu)$ is a measure on Y , provided $\mu \in \$X$. This is fairly straightforward, let's have a look:

1. $\$(f)(\mu)(\emptyset) = \mu(f^{-1}[\emptyset]) = \mu(\emptyset) = 0$, and $\$(f)(\mu)(Y) = \mu(f^{-1}[Y]) = \mu(X) = 1$.
2. Let A and B be disjoint measurable subsets of Y , then $f^{-1}[A]$ and $f^{-1}[B]$ are disjoint as well, thus

$$\begin{aligned} \$(f)(\mu)(A \cup B) &= \mu(f^{-1}[A \cup B]) \\ &= \mu(f^{-1}[A] \cup f^{-1}[B]) \\ &= \mu(f^{-1}[A]) + \mu(f^{-1}[B]) \\ &= \$(f)(\mu)(A) + \$(f)(\mu)(B). \end{aligned}$$

3. If $(B_n)_{n \in \mathbb{N}}$ is an increasing sequence of measurable sets in Y with $B := \bigcup_{n \in \mathbb{N}} B_n$, then $(f^{-1}[B_n])_{n \in \mathbb{N}}$ is an increasing sequence of measurable subsets of X , and $f^{-1}[B]$ equals $\bigcup_{n \in \mathbb{N}} f^{-1}[B_n]$, thus

$$\$(f)(\mu)(B) = \mu(f^{-1}[B]) = \sup_{n \in \mathbb{N}} \mu(f^{-1}[B_n]) = \sup_{n \in \mathbb{N}} \$(f)(\mu)(B_n).$$

1. We establish measurability by showing that the inverse image of a generator to the * - σ -algebra in Y is a * - σ -measurable subset of $\$(X)$. Then the assertion will follow from Lemma 3.5. In fact, let $B \subseteq Y$ be measurable, then we claim that

$$(\$f)^{-1}[\beta_Y(B, q)] = \beta_X(f^{-1}[B], q)$$

holds. This is so because

$$\begin{aligned} \mu \in (\$f)^{-1}[\beta_Y(B, q)] &\Leftrightarrow \$(f)(\mu) \in \beta_Y(B, q) \\ &\Leftrightarrow \$(f)(\mu)(B) \geq q \\ &\Leftrightarrow \mu(f^{-1}[B]) \geq q \\ &\Leftrightarrow \mu \in \beta_X(f^{-1}[B], q). \end{aligned}$$

\dashv

This yields as an immediate consequence

Proposition 3.12 *$\$$ is an endofunctor on the category of measurable spaces. \dashv*

4 The Giry Monad

We want to determine the coalgebras for $\$$. Given a measurable space X , a $\$$ -coalgebra (X, K) is a measurable map $K : X \rightarrow \$X$. It will be necessary to proceed a bit more general, and to characterize measurable maps $X \rightarrow \$ (Y)$ first.

Example 4.1 Let X and Y be measurable spaces, and $K : X \rightarrow \$ (Y)$ be a measurable map (remember: $\$ (Y)$ carries the \ast - σ -algebra). Thus

1. $K(x)$ is for every $x \in X$ a probability measure on (the measurable subsets of) Y .
2. Since $K^{-1}[\beta_Y(B, q)] = \{x \in X \mid K(x)(B) \geq q\}$, we see that the map $x \mapsto K(x)(B)$ is measurable for any fixed measurable set $B \subseteq Y$.

Conversely, if we know that $x \mapsto K(x)(B)$ is measurable for any fixed measurable set $B \subseteq Y$, and that $K(x)$ is always a probability measure on Y , then the identity

$$K^{-1}[\beta_Y(B, q)] = \{x \in X \mid K(x)(B) \geq q\}$$

shows that K is a measurable map $X \rightarrow \$Y$.

Thus we have in particular identified the *coalgebras* (X, K) for the $\$$ -functor as maps $K : X \times \mathcal{A} \rightarrow [0, 1]$ such that

1. $K(x)$ is a probability measure on X ,
2. $x \mapsto K(x)(A)$ is a measurable map for each measurable set A .

(here \mathcal{A} is the σ -algebra on X). K is also known in probabilistic circles as a *Markov kernel* or a *transition probability*. In terms of transition systems: $K(x)(A)$ is the probability of making a transition from x to an element of the measurable set A . \mathbb{M}

Example 4.2 We identify the morphisms for these coalgebras now. Let (X, K) and (Y, L) be $\$$ -coalgebras. A morphism φ for these coalgebras must be a measurable map $\varphi : X \rightarrow Y$, which is compatible with the coalgebraic structure. This means in our case that

$$\begin{array}{ccc} X & \xrightarrow{\varphi} & Y \\ \downarrow K & & \downarrow L \\ \$X & \xrightarrow{\$ \varphi} & \$Y \end{array}$$

commutes. Thus

$$L(\varphi(x))(B) = (L \circ \varphi)(x)(B) = (\$ (\varphi) \circ K)(x)(B) = \$ (\varphi) (K(x))(B) = K(x)(\varphi^{-1}[B]),$$

meaning that the probability of hitting an element of B from $\varphi(x)$ is the same as hitting an element of $\varphi^{-1}[B]$ from x . \mathbb{M}

Coalgebras
for $\$$

For proceeding further, we need the integral of a bounded measurable function. Having this at our disposal, we can investigate the Giry monad and put it into context with other known monads by identifying common properties.

Fix for the moment a measurable space X with σ -algebra \mathcal{A} .

Definition 4.3 Denote by $\mathcal{F}(X, \mathcal{A}) = \mathcal{F}(X)$ the set of all bounded measurable functions on X .

The algebraic structure of $\mathcal{F}(X)$ is easily identified.

Lemma 4.4 $\mathcal{F}(X)$ is a real vector space with $\chi_A \in \mathcal{F}(X)$ iff $A \in \mathcal{A}$.

Proof 0. We know already from Example 3.7 that χ_A constitutes a measurable function iff the set A is measurable.

1. It is sufficient to show that $\mathcal{F}(X)$ is closed under addition and under scalar multiplication. The latter property is fairly easily established through Lemma 3.6, so let's try our hand on the sum. We have for $r \in \mathbb{R}$ and $f, g \in \mathcal{F}(X)$

$$\begin{aligned} \{x \in X \mid f(x) + g(x) < r\} &= \bigcup_{q \in \mathbb{Q}, q < r} \{x \in X \mid f(x) + g(x) < q\} \\ &= \bigcup_{q \in \mathbb{Q}, q < r} \left(\bigcup_{a_1, a_2 \in \mathbb{Q}, a_1 + a_2 \leq q} (\{x \mid f(x) < a_1\} \cap \{x \mid g(x) < a_2\}) \right) \end{aligned}$$

Because both $\{x \mid f(x) < a_1\}$ and $\{x \mid g(x) < a_2\}$ are measurable sets, it follows that $\{x \in X \mid f(x) + g(x) < r\}$ is a measurable set, since \mathbb{Q} is countable. The assertion now follows from Lemma 3.6. \dashv

It is also not difficult to show with the available tools that $\lim_{n \rightarrow \infty} f_n$ defines a member of $\mathcal{F}(X)$, provided $(f_n)_{n \in \mathbb{N}} \subseteq \mathcal{F}(X)$ such that $|f_n(x)| \leq B$ for all $n \in \mathbb{N}$ and $x \in X$, where $B \in \mathbb{R}$ (the latter condition is necessary for making the limit a bounded function).

This permits us to define the integral of a bounded measurable function. The elaborate process is somewhat technical and drawn out in great detail in [9, Section 4.8]; we restrict ourselves to presenting the result.

Proposition 4.5 Let $\mu \in \mathcal{S}(X)$ be a probability measure on X . There exists a unique map $\Phi : \mathcal{F}(X) \rightarrow \mathbb{R}$ with these properties:

1. $\Phi(a \cdot f + b \cdot g) = a \cdot \Phi(f) + b \cdot \Phi(g)$, whenever $a, b \in \mathbb{R}$ and $f, g \in \mathcal{F}(X)$ (linearity).
2. $\Phi(f) \geq 0$, provided $f \geq 0$, hence Φ is monotone (positivity).
3. $\Phi(\chi_A) = \mu(A)$ for all measurable sets $A \subseteq X$ (extension).
4. If $(f_n)_{n \in \mathbb{N}}$ is a sequence of bounded measurable functions such that the limit $\lim_{n \rightarrow \infty} f_n$ is bounded, then

$$\Phi\left(\lim_{n \rightarrow \infty} f_n\right) = \lim_{n \rightarrow \infty} \Phi(f_n).$$

This is usually referred to as continuity.

\dashv

Notation: $\Phi(f)$ is written traditionally as $\int_X f \, d\mu$, or as $\int_X f(x) \, \mu(dx)$, if we want to emphasize the integration variable. We write $\int_A f \, d\mu$ for $\int_X f \cdot \chi_A \, d\mu$. It is called the *integral of f with respect to μ* .

$$\int_X f \, d\mu$$

Example 4.6 Let $K : X \rightarrow \mathcal{S}(Y)$ be a measurable map for the measurable spaces X and Y . Define for $\mu \in \mathcal{S}(X)$

$$K^*(\mu)(B) := \int_X K(x)(B) \, \mu(dx)$$

for $B \subseteq Y$ measurable. Then

$$K^*(\mu)$$

1. $K^*(\mu) \in \mathcal{S}Y$,
2. $\mu \mapsto K^*(\mu)$ is a measurable map $\mathcal{S}(X) \rightarrow \mathcal{S}(Y)$ with respect to the $*$ - σ -algebras on $\mathcal{S}(X)$ resp. $\mathcal{S}(Y)$.

We establish only the first property. The second one is not particularly difficult, but a bit more time consuming to establish, so I refer you to [9, Example 2.4.8].

Let us have a look at the properties of a measure:

1. $K^*(\mu)(\emptyset) = \int_X 0 \, d\mu = 0$ and $K^*(\mu)(Y) = \int_X 1 \, d\mu = \int_X \chi_X \, d\mu = \mu(X) = 1$.
2. Let A and B are disjoint measurable sets in Y , then

$$\begin{aligned} K^*(\mu)(A \cup B) &= \int_X K(x)(A \cup B) \, \mu(dx) \\ &= \int_X (K(x)(A) + K(x)(B)) \, \mu(dx) && \text{(each } K(x) \text{ is a measure)} \\ &= \int_X K(x)(A) \, \mu(dx) + \int_X K(x)(B) \, \mu(dx) && \text{(additivity of the integral)} \\ &= K^*(\mu)(A) + K^*(\mu)(B). \end{aligned}$$

3. Assume that $B_1 \subseteq B_2 \subseteq \dots \subseteq B_n \subseteq \dots$ is an increasing sequence of measurable sets in Y with $B := \bigcup_{n \in \mathbb{N}} B_n$, then

$$K(x)(B) = \lim_{n \rightarrow \infty} K(x)(B_n)$$

for all $x \in X$, since each $K(x)$ is a measure, and $0 \leq K(x)(B) \leq 1$ for all $x \in X$, thus we obtain from continuity

$$\begin{aligned} K^*(\mu)(B) &= \int_X K(x)(B) \, \mu(dx) \\ &= \int_X \lim_{n \rightarrow \infty} K(x)(B_n) \, \mu(dx) \\ &= \lim_{n \rightarrow \infty} \int_X K(x)(B_n) \, \mu(dx) \\ &= \lim_{n \rightarrow \infty} K^*(\mu)(B_n). \end{aligned}$$

The next lemma shows an important technique for working with integrals, perceived as extensions of measures; we will need this property badly when we are discussing the Giry monad. The question arises naturally how to integrate with respect to the measure $K^*(\mu)$, so we will try to piece the integration together from the integrals wrt the measures $K(x)$ for every $x \in X$, and from the integral wrt μ .

Lemma 4.7 *Let X, Y, K, μ as above, then we have for all $f \in \mathcal{F}(Y)$*

1. $x \mapsto \int_Y f(y) K(x)(dy)$ defines a measurable and bounded function on X .
2. $\int_X f(x) K^*(\mu)(dx) = \int_X \left(\int_Y f(y) K(x)(dy) \right) \mu(dx)$.

Proof 0. This is established very similar to the principle of good sets (see the proof of Lemma 3.5). Put

$$\mathcal{E} := \{f \in \mathcal{F}(Y) \mid \text{the assertions are true for } f\}.$$

Then clearly \mathcal{E} is a vector space over \mathbb{R} . We show that $\chi_B \in \mathcal{E}$ for $B \subseteq Y$ measurable; since \mathcal{E} is closed under bounded limits, the assertion follows from the observation that linear combinations of indicator functions are dense in $\mathcal{F}(X)$. The most complicated thing is to show that $\chi_B \in \mathcal{E}$, which we will do now.

1. We claim that both assertions are true for $f = \chi_B$, if $B \subseteq Y$ is a measurable set. This is so because in this case we have for the first claim

$$x \mapsto \int_Y f(y) K(x)(dy) = \int_Y \chi_B(y) K(x)(dy) = K(x)(B)$$

and $x \mapsto K(x)(B)$ defines a measurable function by definition of the $*$ - σ -algebra. From this we obtain

$$\begin{aligned} \int_X \left(\int_Y \chi_B(y) K(x)(dy) \right) \mu(dx) &= \int_X (K(x)(B)) \mu(dx) \\ &= K^*(\mu)(B) && \text{(inner integral)} \\ &= \int_X \chi_B(y) K^*(\mu)(dy) && \text{(extension property)} \end{aligned}$$

2. If $f = \sum_{i=1}^n \alpha_i \cdot \chi_{B_i}$ is a step function with measurable sets B_1, \dots, B_n , the assertion follows from the first part through the additivity of the integral.
3. The measurable step functions are dense in $\mathcal{F}(X)$ with respect to pointwise convergence, so the assertion follows from the second part, and from continuity of the integral. \dashv

This provides us with an amazing consequence.

Proposition 4.8 *Let X, Y and Z be measurable spaces, $K : X \rightarrow \mathcal{S}(Y)$ and $L : Y \rightarrow \mathcal{S}(Z)$ be measurable maps. Then $K^* \circ L^* = (K^* \circ L)^*$.*

Proof 0. The proof is essentially a special case of Lemma 4.7 (although it does not look like it), making use of the fact that $y \mapsto K(y)(C)$ is a measurable map, and that $L(y)$ is a measure for each $y \in Y$ and each $C \subseteq Z$ measurable.

1. Let $\mu \in \mathcal{M}X$, and $C \subseteq Z$ be measurable, then we have

$$\begin{aligned}
 (L^* \circ K^*)(\mu)(C) &= L^*(K^*(\mu))(C) \\
 &= \int_Y L(y)(C) K^*(\mu)(dy) && \text{(definition of } -^*) \\
 &= \int_X \left(\int_Y L(y)(C) K(x)(dy) \right) \mu(dx) && \text{(apply Lemma 4.7)} \\
 &= \int_X L^*(K(x))(C) \mu(dx) && \text{(note that } (\dots) = L^*(K(x))(C)) \\
 &= (L^* \circ K)^*(\mu)(C) && \text{(definition of } -^*)
 \end{aligned}$$

⊢

Example 4.9 Lift $f : X \rightarrow \mathbb{2}^X$ to $f^* : \mathbb{2}^X \rightarrow \mathbb{2}^X$ upon setting

$$f^*(A) := \bigcup_{x \in A} f(x).$$

Then an easy computation which the reader is invited to perform shows that $f^* \circ g^* = (f^* \circ g)^*$ holds. \mathcal{M}

Example 4.10 Lift $f : X \rightarrow \mathbb{E}(X)$ to $f^* : \mathbb{E}(X) \rightarrow \mathbb{E}(X)$ upon setting

$$f^*(\mathcal{C}) := \{B \subseteq X \mid \{x \mid B \in f(x)\} \in \mathcal{C}\}.$$

Then an easy computation with a similar scope shows that $f^* \circ g^* = (f^* \circ g)^*$ holds. \mathcal{M}

This is certainly not such a strange coincidence.

Definition 4.11 Let \mathbf{C} be a category, \mathbf{T} be a map which maps the objects in \mathbf{C} to objects in \mathbf{C} . Assume that we have a map $-^*$ which maps morphisms $f : x \rightarrow \mathbf{T}y$ to morphisms $f^* : \mathbf{T}x \rightarrow \mathbf{T}y$ (called lifting), and a morphism $\eta_x : x \rightarrow \mathbf{T}x$ (called embedding) for each object x in \mathbf{C} . Then $(\mathbf{T}, -^*, \eta)$ is called a monad iff these conditions hold (here x, y, z are objects in \mathbf{C}):

Monad

1. $\eta_x^* = \text{id}_{\mathbf{T}x}$.
2. $f^* \circ \eta_x = f$, whenever $f : x \rightarrow \mathbf{T}y$.
3. $g^* \circ f^* = (g^* \circ f)^*$, whenever $f : x \rightarrow \mathbf{T}y$, $g : y \rightarrow \mathbf{T}z$.

What we call a monad here is usually called a *Kleisli triple* in the literature. By Manes' Theorem [9, Theorem 2.4.4], Kleisli tripels and monads are equivalent. Introducing monads in this way has the advantage of not having to introduce natural transformations and the slightly complicated diagrams associated with it.

Note that \mathbf{T} is only assumed to map objects to objects, but the laws of a monad permit defining it on morphisms as well.

Lemma 4.12 Let $(\mathbf{T}, -^*, \eta)$ be a monad over \mathbf{C} . Then \mathbf{T} can be extended to an endofunctor on \mathbf{C} .

Proof Define $\mathbf{T}f := (\eta_y \circ f)^*$ for $f : x \rightarrow y$. Then $\mathbf{T}f : \mathbf{T}x \rightarrow \mathbf{T}y$ is a morphism in \mathbf{C} with

1. $\mathbf{T}\text{id}_x = \eta_x^* = \text{id}_{\mathbf{T}x}$.
2. Assume $f : x \rightarrow y$ and $g : y \rightarrow z$, then we have

$$\begin{aligned}
 (\mathbf{T}g) \circ (\mathbf{T}f) &= (\eta_z \circ g)^* \circ (\eta_y \circ f)^* && \text{(definition)} \\
 &\stackrel{(\ddagger)}{=} ((\eta_z \circ g)^* \circ \eta_y \circ f)^* \\
 &= (\eta_z \circ g \circ f)^* && \text{(from 3. in Definition 4.11)} \\
 &= \mathbf{T}(g \circ f).
 \end{aligned}$$

Equation (\ddagger) uses the interplay of η and the $-^*$ -operation, see property 2. in Definition 4.11.

⊥

We are now in proud possession of the following monads:

1. Power set monad $\mathbf{2}^-$, f^* according to Example 4.9, $\eta_X(\mathbf{a}) := \{\mathbf{a}\}$ for $\mathbf{a} \in X$.
2. Upper closed monad \mathbf{E} , f^* according to Example 4.10, $\eta_X(\mathbf{a}) := \{A \subseteq X \mid \mathbf{a} \in A\}$ for $\mathbf{a} \in X$.
3. Probability monad $\mathbf{\$}$, K^* according to Example 4.6, η is given through the Dirac kernel from Example 3.9.

Some
monads

If you want to try your hand at other monads, try these:

Some other
monads

1. The *ultra filter monad* over the category of sets. Let $\mathbf{U}(X)$ be all ultrafilters over set X , and define $\mathbf{U}(f) : \mathbf{U}(X) \rightarrow \mathbf{U}(Y)$ for a given map $f : X \rightarrow Y$ verbatim as for upper closed sets in Example 2.8, replacing the argument to $\mathbf{U}(f)$ by an ultrafilter (it has to be shown that $\mathbf{U}(f)(\mathcal{C}) \in \mathbf{U}(Y)$; this requires some thought). Define the embedding $X \rightarrow \mathbf{U}(X)$ as in the case of the upper closed subsets.
2. The *discrete probability monad* over the category of sets. Define

$$\mathbf{D}(X) := \{p : X \rightarrow [0, 1] \mid p \text{ has countable support, and } \sum_{x \in X} p(x) = 1\},$$

where the support of a map $p : X \rightarrow [0, 1]$ is defined as $\{x \in X \mid p(x) \neq 0\}$ (hence the sum is defined). Let $f : X \rightarrow Y$ be a map, $p \in \mathbf{D}(X)$, define

$$\mathbf{D}(f)(p)(y) := \sum_{f(x)=y} p(x).$$

Then show that $\mathbf{D}(f)(p) \in \mathbf{D}(Y)$. The embedding is defined as the Dirac kernel from Example 3.9.

We identify in what follows a monad with its functor, so that things are a bit easier to handle. If, however, we need the components, we will be explicit about them.

Kleisli
category

Proposition 4.13 *A monad \mathbf{T} over category \mathbf{C} generates a new category $\mathbf{C}_{\mathbf{T}}$ in the following way:*

1. *The objects of $\mathbf{C}_{\mathbf{T}}$ are the objects of \mathbf{C} ,*
2. *A $\mathbf{C}_{\mathbf{T}}$ -morphism $f : x \rightsquigarrow y$ in the new category is a \mathbf{C} -morphism $f : x \rightarrow \mathbf{T}y$ in \mathbf{C} .*
3. *The identity for a in $\mathbf{C}_{\mathbf{T}}$ is $\eta_a : a \rightarrow \mathbf{T}a$.*
4. *The composition $g * f$ of $f : x \rightsquigarrow y$ and $g : y \rightsquigarrow z$ is defined through $g * f := g^* \circ f$.*

$f : x \rightsquigarrow y$

This category is called the Kleisli category associated with \mathbf{T} (and \mathbf{C} , of course).

Proof We have to show that the laws of a category are satisfied, hence in particular that Kleisli composition is associative. In fact, we have

$$\begin{aligned}
 (h * g) * f &= (h * g)^* \circ f \\
 &= (h^* \circ g)^* \circ f && \text{(definition of } h * g\text{)} \\
 &= h^* \circ g^* \circ f && \text{(property 3. in a monad)} \\
 &= h^* \circ (g * f) && \text{(definition of } g * f\text{)} \\
 &= h * (f * g)
 \end{aligned}$$

The laws for the identity are easily checked from the properties of $-^*$. Thus $\mathbf{C}_{\mathbf{T}}$ is indeed a category. \dashv

Example 4.14 The Kleisli morphisms for the power set monad are exactly the relations, and we have for $R : X \rightsquigarrow Y$ and $S : Y \rightsquigarrow Z$ that

$$z \in (R * S)(x) \Leftrightarrow z \in S(y) \text{ for some } y \in R(x).$$

This is immediate. M

Example 4.15 The Kleisli morphisms for the Giry monad are exactly the stochastic relations, a.k.a. transition probabilities. Let $K : X \rightsquigarrow Y$ and $L : Y \rightsquigarrow Z$ be stochastic relations, then we have

$$(L * K)(x)(C) = \int_Y L(y)(C) K(x)(dy),$$

when $x \in X$ and $C \subseteq Z$ is a measurable set. This follows immediately from the definition of $-^*$ for this monad. M

5 Playing Around with Morphisms

Measurable spaces form a category under measurable maps. In fact, given a measurable space (X, \mathcal{A}) ,

1. we can find for a map $f : Z \rightarrow X$ a smallest σ -algebra \mathcal{C} on Z which renders f a (Z, \mathcal{C}) -(X, \mathcal{A})-measurable map. Take simply

$$\mathcal{C} := \{f^{-1}[A] \mid A \in \mathcal{A}\}.$$

\mathcal{C} is called the *initial σ -algebra* with respect to f and \mathcal{A} .

Initial

2. we can find for a map $g : X \rightarrow Y$ a largest σ -algebra \mathcal{B} on Y such that g is \mathcal{A} -(Y, \mathcal{B})-measurable. Take simply

$$\mathcal{B} := \{B \subseteq Y \mid f^{-1}[B] \in \mathcal{A}\}.$$

\mathcal{B} is called the *final σ -algebra* with respect to g and \mathcal{A} .

Final

Both constructions extend easily to families of maps $f_i : Z_i \rightarrow X$ resp. $g_i : X \rightarrow Y_i$. For example, the \ast - σ -algebra on $\$(X)$ is the initial σ -algebra with respect to the family $\{\text{ev}_A \mid A \in \mathcal{A}\}$, with $\text{ev}_A : \mu \mapsto \mu(A)$, and the *product- σ -algebra* $\mathcal{A} \otimes \mathcal{B}$ on the Cartesian product $X \times Y$ is the initial σ -algebra on $X \times Y$ with respect to the projections π_X and π_Y .

 $\mathcal{A} \otimes \mathcal{B}$

Given an equivalence relation τ on X , define \mathcal{A}/τ as the final σ -algebra on the set X/τ of equivalence classes with respect to the factor map $\rho_\tau : x \mapsto [x]_\tau$. We assume that this space is always equipped with this σ -algebra.

Equivalence relation τ defines also a σ -algebra on X , the σ -algebra of τ -invariant (*measurable sets*), upon setting

 Σ_τ

$$\Sigma_\tau := \Sigma_{\tau, \mathcal{A}} := \{A \in \mathcal{A} \mid A \text{ is } \tau\text{-invariant}\}$$

(recall that set A is τ -invariant iff it is the union of τ -classes, or, equivalently, iff $x \in A$ and $x \tau x'$ together imply $x' \in A$). Look at this equivalence

$$A \in \Sigma_\tau \Leftrightarrow \rho_\tau[A] \in \mathcal{A}/\tau.$$

Does it always hold? In fact, this is true, and it hinges on the equality $\rho_\tau^{-1}[\rho_\tau[A]] = A$ for τ -invariant $A \subseteq X$ (\subseteq : If $\rho_\tau(x) \in \rho_\tau[A]$, there exists $x' \in A$ with $x \rho x'$, hence $x \in A$; \supseteq : is trivial). An equivalent formulation is evidently

$$\mathcal{A}/\tau = \{A \subseteq X/\tau \mid \rho_\tau^{-1}[A] \in \Sigma_\tau\}.$$

Σ_τ is a fairly important σ -algebra, as we will see. Occasionally one considers as an equivalence relation the *kernel* $\ker(f)$ of a measurable map $f : X \rightarrow Y$. It is defined as

Kernel,
 $\ker(f)$

$$\ker(f) := \{\langle x, x' \rangle \mid f(x) = f(x')\}.$$

We are now in a position to define congruences for stochastic relations on a measurable space X . Remember that a stochastic relation $K : X \rightsquigarrow X$ is a coalgebra (X, K) for the Giry functor.

Definition 5.1 Let $K : X \rightsquigarrow X$ be a stochastic relation. An equivalence relation τ is called a congruence for K iff there exists a stochastic relation $K_\tau : X/\tau \rightsquigarrow X/\tau$ such that $\rho_\tau : K \rightarrow K_\tau$ is a morphism.

Let us see what it means that τ is a congruence for K . Since we are dealing with coalgebras here, this means that this diagram commutes

$$\begin{array}{ccc} X & \xrightarrow{\rho_\tau} & X/\tau \\ K \downarrow & & \downarrow K_\tau \\ \$ (X) & \xrightarrow{\$(\rho_\tau)} & \$ (X/\tau) \end{array}$$

Hence we have for $x \in X$ and $A \subseteq X/\tau$ measurable this equality

$$K_\tau([x]_\tau)(A) = (K_\tau \circ \rho_\tau)(x)(A) = (\$(\rho_\tau) \circ K(x))(A) = K(x)(\rho_\tau^{-1}[A]).$$

This means that the behavior of $K(x)$ on the σ -algebra Σ_τ determines the behavior of $K_\tau([x]_\tau)$ completely. This is intuitively somewhat satisfying: if τ cannot distinguish between x and x' , then $K(x)(A)$ should be the same as $K(x')(A)$ for all A the elements of which τ cannot tell apart (actually, this is how a congruence was first defined for K).

In universal algebra there is a strong connection between the kernels of morphisms and congruences, actually, e.g., in Abelian groups, the kernel of a morphism is a congruence, and vice versa. In general, additional conditions are necessary. A measurable map $f : X \rightarrow Y$ is called *strong* iff f is surjective so that Y carries the final σ -algebra with respect to f ; note that being strong is an intrinsic property of f and is independent of any $\$$ -coalgebra.

Strong

Proposition 5.2 Let (X, K) and (Y, L) be $\$$ -coalgebras, and $f : (X, K) \rightarrow (Y, L)$ is a strong morphism. Then $\ker(f)$ is a congruence. Conversely, if τ is a congruence for (X, K) , then ρ_τ is a strong morphism.

Proof The assertion about ρ_τ is trivial from the construction, and since $\ker(\rho_\tau) = \tau$. The converse follows from some observations on general coalgebras based on sets, and a characterization of Σ_f for strong f in [9, Section 2.6.2]. \dashv

We want to define subsystems for a $\$$ -coalgebra (X, K) , $X = (X, \mathcal{A})$ being a measurable space again. Subsystems are determined through a sub- σ -algebra $\mathcal{B} \subseteq \mathcal{A}$ and through a transition law, say, L . Note that the identity $i_X : X \rightarrow X$ is \mathcal{A} - \mathcal{B} -measurable iff $\mathcal{B} \subseteq \mathcal{A}$. This time we have to make the σ -algebra explicit.

Definition 5.3 $((X, \mathcal{B}), L)$ is a subsystem of $((X, \mathcal{A}), K)$ iff the identity is a morphism $i_X : ((X, \mathcal{A}), K) \rightarrow ((X, \mathcal{B}), L)$.

Subsystem

Again, we have this diagram, which commutes for a subsystem:

$$\begin{array}{ccc} (X, \mathcal{A}) & \xrightarrow{i_X} & (X, \mathcal{B}) \\ \downarrow \kappa & & \downarrow \mathsf{L} \\ \$(X, \mathcal{A}) & \xrightarrow{\$(i_X)} & \$(X, \mathcal{B}) \end{array}$$

Hence $K(x)(B) = L(x)(B)$ for all $x \in X$, and all $B \in \mathcal{B}$, which may be interpreted either that $K(x)$ is an extension to $L(x)$ or that $L(x)$ is the restriction of $K(x)$, depending on the situation at hand. Sometimes a subsystem is called a *state bisimulation*, but I think that this is an unfortunate name, because bisimilarity as a means of comparing the expressivity of systems through a mediator is nowhere to be seen. A subsystem will be identified it through its σ -algebra \mathcal{B} ; the coalgebra is then defined through the restriction to \mathcal{B} .

It is immediate that a congruence τ defines a subsystem with Σ_τ as the defining σ -algebra.

6 Bisimulations

The notion of bisimilarity is fundamental for the application of coalgebras to system modelling. Bisimilar coalgebras behave in a similar fashion, witnessed by a mediating system.

Definition 6.1 *Let \mathbf{F} be an endofunctor on a category \mathbf{C} . The \mathbf{F} -coalgebras (a, f) and (b, g) are said to be bisimilar iff there exists a coalgebra (m, v) and coalgebra morphisms*

Bisimilar,
mediating

$$(a, f) \longleftarrow (m, v) \longrightarrow (b, g).$$

The coalgebra (m, v) is called mediating.

Thus we obtain this characteristic diagram with ℓ and r as the corresponding morphisms.

$$\begin{array}{ccccc} a & \xleftarrow{\ell} & m & \xrightarrow{r} & b \\ f \downarrow & & \downarrow v & & \downarrow g \\ \mathbf{F}a & \xleftarrow{\mathbf{F}\ell} & \mathbf{F}m & \xrightarrow{\mathbf{F}r} & \mathbf{F}b \end{array}$$

This gives us $f \circ \ell = (\mathbf{F}\ell) \circ v$ and together with $g \circ r = (\mathbf{F}r) \circ v$. It is easy to see why (M, m) is called mediating.

Bisimilarity was originally investigated when concurrent systems became of interest [14]. The original formulation, however, was not coalgebraic but rather relational. Here it is (for the sake of easier reading, we use arrows rather than relations or maps into the respective power set):

Definition 6.2 *Let (S, \rightsquigarrow_S) and (T, \rightsquigarrow_T) be transition systems. Then $B \subseteq S \times T$ is called a bisimulation iff for all $\langle s, t \rangle \in B$ these conditions are satisfied:*

Bisimulation,
2-

1. if $s \rightsquigarrow_S s'$, then there is a $t' \in T$ such that $t \rightsquigarrow_T t'$ and $\langle s', t' \rangle \in B$,

2. if $t \rightsquigarrow_T t'$, then there is a $s' \in S$ such that $s \rightsquigarrow_S s'$ and $\langle s', t' \rangle \in B$.

Hence a bisimulation simulates transitions in one system through the other one. On first sight, these notions of bisimilarity are not related to each other. Recall that transition systems are coalgebras for the power set functor $\mathbb{2}^-$. This is the connection:

Theorem 6.3 *Given the transition systems (S, \rightsquigarrow_S) and (T, \rightsquigarrow_T) with the associated $\mathbb{2}^-$ -coalgebras (S, f) and (T, g) , then these statements are equivalent for $B \subseteq S \times T$:*

Aczel's
Theorem

1. B is a bisimulation.
2. There exists a $\mathbb{2}^-$ -coalgebra structure h on B such that $(S, f) \longleftarrow (B, h) \longrightarrow (T, g)$ with the projections as morphisms is mediating.

Proof That $(S, f) \xleftarrow{\pi_S} (B, h) \xrightarrow{\pi_T} (T, g)$ is mediating follows from commutativity of this diagram.

$$\begin{array}{ccccc}
 S & \xleftarrow{\pi_S} & B & \xrightarrow{\pi_T} & T \\
 f \downarrow & & h \downarrow & & g \downarrow \\
 \mathbb{2}^-(S) & \xleftarrow{\mathbb{2}^-(\pi_S)} & \mathbb{2}^-(B) & \xrightarrow{\mathbb{2}^-(\pi_T)} & \mathbb{2}^-(T)
 \end{array}$$

“1 \Rightarrow 2”: We have to construct a map $h : B \rightarrow \mathbb{2}^-(B)$ such that $f(\pi_S(s, t)) = \mathbb{2}^-(\pi_S)(h(s, t))$ and $f(\pi_T(s, t)) = \mathbb{2}^-(\pi_T)(h(s, t))$ for all $\langle s, t \rangle \in B$. The choice is somewhat obvious: put for $\langle s, t \rangle \in B$

$$h(s, t) := \{\langle s', t' \rangle \in B \mid s \rightsquigarrow_S s', t \rightsquigarrow_T t'\}.$$

Thus $h : B \rightarrow \mathbb{2}^-(B)$ is a map, hence (B, h) is a $\mathbb{2}^-$ -coalgebra.

Now fix $\langle s, t \rangle \in B$, then we claim that $f(s) = \mathbb{2}^-(\pi_S)(h(s, t))$.

“ \subseteq ”: Let $s' \in f(s)$, hence $s \rightsquigarrow_S s'$, thus there exists t' with $\langle s', t' \rangle \in B$ such that $t \rightsquigarrow_T t'$, hence

$$s' \in \{\pi_S(s_0, t_0) \mid \langle s_0, t_0 \rangle \in h(s, t)\} = \{s_0 \mid \langle s_0, t_0 \rangle \in h(s, t) \text{ for some } t_0\} = \mathbb{2}^-(\pi_S)(h(s, t)).$$

“ \supseteq ”: If $s' \in \mathbb{2}^-(\pi_S)(h(s, t))$, then in particular $s \rightsquigarrow_S s'$, thus $s' \in f(s)$.

Thus we have shown that $\mathbb{2}^-(\pi_S)(h(s, t)) = f(s) = f(\pi_S(s, t))$. One shows $\mathbb{2}^-(\pi_T)(h(s, t)) = g(t) = f(\pi_T(s, t))$ in exactly the same way. We have constructed h such that (B, h) is a $\mathbb{2}^-$ -coalgebra, and such that the diagrams above commute.

“2 \Rightarrow 1”: Assume that h exists with the properties described in the assertion, then we have to show that B is a bisimulation. Now let $\langle s, t \rangle \in B$ and $s \rightsquigarrow_S s'$, hence $s' \in f(s) = f(\pi_S(s, t)) = \mathbb{2}^-(\pi_S)(h(s, t))$. Thus there exists t' with $\langle s', t' \rangle \in h(s, t) \subseteq B$, and hence $\langle s', t' \rangle \in B$. We claim that $t \rightsquigarrow_T t'$, which is tantamount to saying $t' \in g(t)$. But $g(t) = \mathbb{2}^-(\pi_T)(h(s, t))$, and $\langle s', t' \rangle \in h(s, t)$, hence $t' \in \mathbb{2}^-(\pi_T)(h(s, t)) = g(t)$. This establishes $t \rightsquigarrow_T t'$. A similar argument finds s' with $s \rightsquigarrow_S s'$ with $\langle s', t' \rangle \in B$ in case $t \rightsquigarrow_T t'$.

This completes the proof. \dashv

Thus we may use bisimulations for transition systems as relations and bisimulations as coalgebras interchangeably, and this characterization suggests a definition in purely coalgebraic terms for

those cases in which a set-theoretic relation is not available or not adequate. The connection to $\mathbf{2}^-$ -coalgebra morphisms and bisimulations is further strengthened by investigating the graph of a morphism (recall that the graph of a map $r : S \rightarrow T$ is the relation $\text{graph}(r) := \{\langle s, r(s) \rangle \mid s \in S\}$).

$\text{graph}(r)$

Proposition 6.4 *Given coalgebras (S, f) and (T, g) for the power set functor $\mathbf{2}^-$, $r : (S, f) \rightarrow (T, g)$ is a morphism iff $\text{graph}(r)$ is a bisimulation for (S, f) and (T, g) .*

Proof 1. Assume that $r : (S, f) \rightarrow (T, g)$ is a morphism, so that $g \circ r = \mathbf{2}^-(r) \circ f$. Now define

$$h(s, t) := \{\langle s', r(s') \rangle \mid s' \in f(s)\} \subseteq \text{graph}(r)$$

for $\langle s, t \rangle \in \text{graph}(r)$. Then $g(\pi_T(s, t)) = g(t) = \mathbf{2}^-(\pi_T)(h(s, t))$ for $t = r(s)$.

“ \subseteq ”: If $t' \in g(t)$ for $t = r(s)$, then

$$\begin{aligned} t' \in g(r(s)) &= \mathbf{2}^-(r)(f(s)) &&= \{r(s') \mid s' \in f(s)\} \\ &= \mathbf{2}^-(\pi_T)(\{\langle s', r(s') \rangle \mid s' \in f(s)\}) &&= \mathbf{2}^-(\pi_T)(h(s, t)) \end{aligned}$$

“ \supseteq ”: If $\langle s', t' \rangle \in h(s, t)$, then $s' \in f(s)$ and $t' = r(s')$, but this implies $t' \in \mathbf{2}^-(r)(f(s)) = g(r(s))$.

Thus $g \circ \pi_T = \mathbf{2}^-(\pi_T) \circ h$. The equation $f \circ \pi_S = \mathbf{2}^-(\pi_S) \circ h$ is established similarly.

Hence we have found a coalgebra structure h on $\text{graph}(r)$ such that

$$(S, f) \xleftarrow{\pi_S} (\text{graph}(r), h) \xrightarrow{\pi_T} (T, g)$$

are coalgebra morphisms, so that $(\text{graph}(r), h)$ is now officially a bisimulation.

2. If, conversely, $(\text{graph}(r), h)$ is a bisimulation with the projections as morphisms, then we have $r = \pi_T \circ \pi_S^{-1}$. Then π_T is a morphism, and π_S^{-1} is a morphism as well (note that we work on the graph of r). So r is a morphism. \dashv

Let us have a look at upper closed sets. There we find a comparable situation. We cannot, however, translate the definition directly, because we do not have access to the transitions proper, but rather to the sets from which the next state may come from. Let (S, f) and (T, g) be \mathbb{E} -coalgebras, and assume that $\langle s, t \rangle \in B$. Assume $X \in f(s)$, then we want to find $Y \in g(t)$ such that, when we take $t' \in Y$, we find a state $s' \in X$ with s' being related via B to s' , and vice versa. Formally:

Definition 6.5 *Let again*

$$\mathbb{E}S := \{V \subseteq \mathbf{2}^-(S) \mid V \text{ is upper closed}\}$$

Bisimulation,

\mathbb{E}

be the endofunctor on **Set** which assigns to set S all upper closed subsets of $\mathbf{2}^-S$. Given \mathbb{E} -coalgebras (S, f) and (T, g) , a subset $B \subseteq S \times T$ is called a bisimulation of (S, f) and (T, g) iff for each $\langle s, t \rangle \in B$

1. for all $X \in f(s)$ there exists $Y \in g(t)$ such that for each $t' \in Y$ there exists $s' \in X$ with $\langle s', t' \rangle \in B$,

2. for all $Y \in g(t)$ there exists $X \in f(s)$ such that for each $s' \in X$ there exists $t' \in Y$ with $\langle s', t' \rangle \in B$.

We have then a comparable characterization of bisimilar coalgebras [7].

Proposition 6.6 *Let (S, f) and (T, g) be coalgebras for \mathbb{E} . Then the following statements are equivalent for $B \subseteq S \times T$ with $\pi_S[B] = S$ and $\pi_T[B] = T$*

1. B is a bisimulation of (S, f) and (T, g) .
2. There exists a coalgebra structure h on B so that the projections $\pi_S : B \rightarrow S, \pi_T : B \rightarrow T$ are morphisms $(S, f) \xleftarrow{\pi_S} (B, h) \xrightarrow{\pi_T} (T, g)$.

Proof “1 \Rightarrow 2”: Define $\langle s, t \rangle \in B$

$$h(s, t) := \{D \subseteq B \mid \pi_S[D] \in f(s) \text{ and } \pi_T[D] \in g(t)\}.$$

Hence $h(s, t) \subseteq 2^B$, and because both $f(s)$ and $g(t)$ are upper closed, so is $h(s, t)$.

Now fix $\langle s, t \rangle \in B$. We show first that $f(s) = \{\pi_S[Z] \mid Z \in h(s, t)\}$. From the definition of $h(s, t)$ it follows that $\pi_S[Z] \in f(s)$ for each $Z \in h(s, t)$. So we have to establish the other inclusion. Let $X \in f(s)$, then $X = \pi_S[\pi_S^{-1}[X]]$, because $\pi_S : B \rightarrow S$ is onto, so it suffices to show that $\pi_S^{-1}[X] \in h(s, t)$, hence that $\pi_T[\pi_S^{-1}[X]] \in g(t)$. Given X there exists $Y \in g(t)$ so that for each $t' \in Y$ there exists $s' \in X$ such that $\langle s', t' \rangle \in B$. Thus $Y = \pi_T[(X \times Y) \cap B]$. But this implies $Y \subseteq \pi_T[\pi_S^{-1}[X]]$, hence $Y \subseteq \pi_T[\pi_S^{-1}[X]] \in g(t)$. One similarly shows that $g(t) = \{\pi_T[Z] \mid Z \in h(s, t)\}$.

In a second step, we show that

$$\{\pi_S[Z] \mid Z \in h(s, t)\} = \{C \mid \pi_S^{-1}[C] \in h(s, t)\}.$$

In fact, if $C = \pi_S[Z]$ for some $Z \in h(s, t)$, then $Z \subseteq \pi_S^{-1}[C] = \pi_S^{-1}[\pi_S[Z]]$, hence $\pi_S^{-1}[C] \in h(s, t)$. If, conversely, $Z := \pi_S^{-1}[C] \in h(s, t)$, then $C = \pi_S[Z]$. Thus we obtain

$$f(s) = \{\pi_S[Z] \mid Z \in h(s, t)\} = \{C \mid \pi_S^{-1}[C] \in h(s, t)\} = (\mathbb{E}\pi_S)(h(s, t))$$

for $\langle s, t \rangle \in B$. Summarizing, this means that $\pi_S : (B, h) \rightarrow (S, f)$ is a morphism. A very similar argumentation shows that $\pi_T : (B, h) \rightarrow (T, g)$ is a morphism as well.

“2 \Rightarrow 1”: Assume, conversely, that the projections are coalgebra morphisms, and let $\langle s, t \rangle \in B$. Given $X \in f(s)$, we know that $X = \pi_S[Z]$ for some $Z \in h(s, t)$. Thus we find for any $t' \in Y$ some $s' \in X$ with $\langle s', t' \rangle \in B$. The symmetric property of a bisimulation is established exactly in the same way. Hence B is a bisimulation for (S, f) and (T, g) . \dashv

We will now turn to bisimulations for stochastic systems. A bisimulation relates two transition systems which are connected through a mediating system. In order to define this for the present context, we extend the crucial notion of morphisms slightly in a straightforward manner; this will be helpful in the discussions to follow.

Definition 6.7 A morphism $m = (f, g) : K_1 \rightarrow K_2$ for stochastic relations $K_i : (X_i, \mathcal{A}_i) \rightsquigarrow (Y_i, \mathcal{B}_i)$

Morphism,
again

($i = 1, 2$) over general measurable spaces is given through the measurable maps $f : X_1 \rightarrow X_2$ and $g : Y_1 \rightarrow Y_2$ such that this diagram of measurable maps commutes

$$\begin{array}{ccc} (X_1, \mathcal{A}_1) & \xrightarrow{f} & (X_2, \mathcal{A}_2) \\ \downarrow \kappa_1 & & \downarrow \kappa_2 \\ \$ (Y_1, \mathcal{B}_1) & \xrightarrow{\$(g)} & \$ (Y_2, \mathcal{B}_2) \end{array}$$

Equivalently, $\kappa_2(f(x_1)) = \$ (g) (\kappa_1(x_1))$, which translates to $\kappa_2(f(x_1))(B) = \kappa_1(x_1)(g^{-1}[B])$ for all $B \in \mathcal{B}_2$.

Bisimilarity,
\$

Definition 6.8 The stochastic relations $\kappa_i : (X_i, \mathcal{A}_i) \rightsquigarrow (Y_i, \mathcal{B}_i)$ ($i = 1, 2$), are called bisimilar iff there exist a stochastic relation $M : (A, \mathcal{X}) \rightsquigarrow (B, \mathcal{Y})$ and surjective morphisms $m_i = (f_i, g_i) : M \rightarrow \kappa_i$ such that the σ -algebra $g_1^{-1}[\mathcal{B}_1] \cap g_2^{-1}[\mathcal{B}_2]$ is nontrivial, i.e., contains not only \emptyset and B . The relation M is called mediating.

The first condition on bisimilarity is in accordance with the general definition of bisimilarity of coalgebras; it requests that m_1 and m_2 form a span of morphisms

$$\kappa_1 \xleftarrow{m_1} M \xrightarrow{m_2} \kappa_2.$$

Hence, the following diagram of measurable maps is supposed to commute with $m_i = (f_i, g_i)$ for $i = 1, 2$

$$\begin{array}{ccccc} (X_1, \mathcal{A}_1) & \xleftarrow{f_1} & (A, \mathcal{X}) & \xrightarrow{f_2} & (X_2, \mathcal{A}_2) \\ \downarrow \kappa_1 & & \downarrow M & & \downarrow \kappa_2 \\ \$ (Y_1, \mathcal{B}_1) & \xleftarrow{\$(g_1)} & \$ (B, \mathcal{Y}) & \xrightarrow{\$(g_2)} & \$ (Y_2, \mathcal{B}_2) \end{array}$$

Thus, for each $a \in A, D \in \mathcal{B}_1, E \in \mathcal{B}_2$ the equalities

$$\begin{aligned} \kappa_1(f_1(a))(D) &= (\$(g_1) \circ M)(a)(D) = M(a)(g_1^{-1}[D]) \\ \kappa_2(f_2(a))(E) &= (\$(g_2) \circ M)(a)(E) = M(a)(g_2^{-1}[E]) \end{aligned}$$

should be satisfied. The second condition, however, is special; it states that we can find an event $C^* \in \mathcal{Y}$ which is common to both κ_1 and κ_2 in the sense that

$$g_1^{-1}[B_1] = C^* = g_2^{-1}[B_2]$$

for some $B_1 \in \mathcal{B}_1$ and $B_2 \in \mathcal{B}_2$ such that both $C^* \neq \emptyset$ and $C^* \neq B$ hold (note that for $C^* = \emptyset$ or $C^* = B$ we can always take the empty and the full set, respectively). Given such a C^* with B_1, B_2 from above we get for each $a \in A$

$$\kappa_1(f_1(a))(B_1) = M(a)(g_1^{-1}[B_1]) = M(a)(C^*) = M(a)(g_2^{-1}[B_2]) = \kappa_2(g_2(a))(B_2);$$

thus the event C^* ties κ_1 and κ_2 together. Loosely speaking, $g_1^{-1}[\mathcal{B}_1] \cap g_2^{-1}[\mathcal{B}_2]$ can be described as the σ -algebra of common events, which is required to be nontrivial.

Note that without the second condition two relations K_1 and K_2 would always be bisimilar: Put $A := X_1 \times X_2$, $B := Y_1 \times Y_2$ and set for $\langle x_1, x_2 \rangle \in A$ as the mediating relation $M(x_1, x_2) := K_1(x_1) \otimes K_2(x_2)$; that is, define M pointwise to be the product measure of K_1 and K_2 . Then the projections will make the diagram commutative. But although this notion of bisimilarity is sometimes suggested, it is way too weak, because bisimulations relate transition systems, and it does not promise particularly interesting insights when two arbitrary systems can be related. It is also clear that using products for mediation does not work for the subprobabilistic case. But the definition above captures the general case as well.

7 Expressivity of Kripke Models

Transition kernels will be used now for interpreting modal logics. Consider this grammar for formulas

$$\varphi ::= \top \mid \varphi_1 \wedge \varphi_2 \mid \Diamond_q \varphi$$

with $q \in \mathbb{Q}, q \geq 0$. Note that the logic is negation free and has on the propositional level only conjunction; this may be motivated by the observation that we work in a Boolean set algebra in which negation is available.

The informal interpretation in a probabilistic transition system is that \top always holds, and that $\Diamond_q \varphi$ holds with probability not smaller than q after a transition in a state in which formula φ holds. Now let $K : X \rightsquigarrow X$ be a transition kernel for the measurable space X , and define inductively

$$\begin{aligned} \llbracket \top \rrbracket_K &:= X \\ \llbracket \varphi_1 \wedge \varphi_2 \rrbracket_K &:= \llbracket \varphi_1 \rrbracket_K \cap \llbracket \varphi_2 \rrbracket_K \\ \llbracket \Diamond_q \varphi \rrbracket_K &:= \{x \in X \mid K(x)(\llbracket \varphi \rrbracket_K) \geq q\} \\ &= K^{-1}[\beta_{\mathcal{A}}(\llbracket \varphi \rrbracket_K, q)] \end{aligned}$$

One shows by induction on the structure of the formula that the sets $\llbracket \varphi \rrbracket_K$ are measurable, since K is a stochastic relation. We say that φ holds in $x \in X$ iff $x \in \llbracket \varphi \rrbracket_K$; this is also written as $K, x \models \varphi$. Note that φ does not hold in x iff $x \in X \setminus \llbracket \varphi \rrbracket_K$, the latter set being measurable. This observation supports the decision to omit negation as an operator.

One usually takes a set of actions and defines modalities $\langle a \rangle_r$ for action a , generalizing \Diamond_r . For the sake of argument, I will stick for the time being to the very simple case of having only one action. The arguments for the general case will be exactly the same, taking into account that one deals with a family of stochastic relations rather than with one relation. I will also do without primitive formulas and introduce them only when I need them; they do not add to the argument's substance right now. A *Kripke model* with state space X and transition law K is just a \mathbb{Q} -coalgebra (X, K) , for the extensions see page 28.

Define for state $x \in X$ its *theory* by

$$\text{Th}_K(x) := \{\varphi \mid x \in \llbracket \varphi \rrbracket_K\} = \{\varphi \mid K, x \models \varphi\}$$

Kripke
model

$\text{Th}_K(x)$

For comparing the expressivity of Kripke models, we use these approaches

Expressivity

- (X, K) is *logically equivalent* to (Y, L) iff

$$\{\text{Th}_K(x) \mid x \in X\} = \{\text{Th}_L(y) \mid y \in Y\},$$

thus iff given a state $x \in X$, there exists a state $y \in Y$ with exactly the same theory, and vice versa.

- (X, K) is *behaviorally equivalent* to (Y, L) iff there exists a Kripke model (Z, M) and surjective morphisms

$$(X, K) \xrightarrow{f} (Z, M) \xleftarrow{g} (Y, L).$$

Model (Z, M) is called *mediating* (and the diagram a *co-span*).

Span,
co-span

- (X, K) is *bisimilar* to (Y, L) iff there exists a Kripke model (Z, M) and surjective morphisms

$$(X, K) \xleftarrow{f} (Z, M) \xrightarrow{g} (Y, L)$$

such that the σ -algebra of common events is not trivial. Model (Z, M) is also called *mediating* (and the diagram a *span*).

We will investigate these notions of expressivity now. Note that there are many variants to bisimilarity, e.g., state bisimulations, but life is difficult enough, so I will not deal with them here.

The key property in this business is

Proposition 7.1 *Let (X, K) and (Y, L) be Kripke models and a morphism $f : (X, K) \rightarrow (Y, L)$. Then*

$$K, x \models \varphi \text{ iff } L, f(x) \models \varphi$$

holds for all $x \in X$ and all formulas φ .

Proof The assertion is equivalent to $\llbracket \varphi \rrbracket_K = f^{-1}[\llbracket \varphi \rrbracket_L]$ for all φ . This is clear for $\varphi = \top$, and if it is true for formulas φ_1 and φ_2 , then it is true also for $\varphi_1 \wedge \varphi_2$. The interesting case is $\varphi_1 = \Diamond_r \varphi$:

$$\begin{aligned} \llbracket \Diamond_r \varphi \rrbracket_K &= \{x \in X \mid K(x)(\llbracket \varphi \rrbracket_K) \geq r\} && \text{(definition)} \\ &= \{x \in X \mid K(x)(f^{-1}[\llbracket \varphi \rrbracket_L]) \geq r\} && \text{(induction hypothesis)} \\ &= \{x \in X \mid L(f(x))(\llbracket \varphi \rrbracket_L) \geq r\} && \text{(f is a morphism)} \\ &= f^{-1}[\llbracket \Diamond_r \varphi \rrbracket_L] \end{aligned}$$

⊢

This provides us with an easy consequence:

Corollary 7.2 *Behaviorally equivalent Kripke models are logically equivalent, so are bisimilar Kripke models.*

The classical Hennessy-Milner Theorem [9, Theorem 2.7.32] for transition systems states that logically equivalent models are behaviorally equivalent, provided the models are image finite. This condition used to prevent the system from becoming too large, but is difficult to model for stochastic Kripke models. Hence we need a condition which is intended restrict the size of the system, so we need a condition for smallness. Here we proceed as follows.

Fix a Kripke model (X, K) . The logic induces an equivalence relation α on X upon setting

$$x \alpha x' \Leftrightarrow [K, x \models \varphi \text{ iff } K, x' \models \varphi] \text{ for all formulas } \varphi.$$

Thus $x \alpha x'$ iff the logic cannot distinguish between x and x' . The discussion above shows that we may factor X , obtaining the factor space X/α . Moreover we know that $K(x)(\llbracket \varphi \rrbracket_K) = K(x')(\llbracket \varphi \rrbracket_K)$, provided $x \alpha x'$ (suppose there exists r with $K(x)(\llbracket \varphi \rrbracket_K) < r \leq K(x')(\llbracket \varphi \rrbracket_K)$, hence $K, x' \models \Diamond_r \varphi$, but $K, x \not\models \Diamond_r \varphi$). This means that α is a congruence for K , the σ -algebra on the factor space being generated by the α -invariant sets $\Sigma_{X,\alpha}$. On the other hand, we have the σ -algebra $\Theta_X := \sigma(\{\llbracket \varphi \rrbracket_K \mid \varphi \text{ is a formula}\})$ which is generated by the validity sets for the formulas (recall that each validity set is measurable). Since each validity set is α -invariant, we have $\Theta_X \subseteq \Sigma_{X,\alpha}$. It may happen, however, that the containment is proper [6, Example 2.6.7].

Definition 7.3 *The Kripke model is said to be small iff $\Sigma_\alpha = \sigma(\{\llbracket \varphi \rrbracket_K \mid \varphi \text{ is a formula}\})$.*

Smallness will permit us to establish an analogon to the Hennessy-Milner Theorem; technically, it says that the σ -algebra on the factor space is determined by the images of the validity sets under the factor map ρ_α :

Lemma 7.4 *If (X, K) is small, then $\mathcal{A}/\alpha = \sigma(\{\rho_\alpha[\llbracket \varphi \rrbracket_K] \mid \varphi \text{ is a formula}\})$. \dashv*

Smallness also ensures that we have a \cap -stable generator for the factor σ -algebra (note that $\rho_\alpha[A \cap B] = \rho_\alpha[A] \cap \rho_\alpha[B]$ if A and B are α -invariant), so that we have — in view of the π - λ -Theorem — a fairly practical handle to deal with the factor space. This will be seen in a moment.

Recall that topological space is called *Polish* iff it is second countable, and its topology can be metrized by a complete metric; examples include the reals \mathbb{R} , $[0, 1]^\mathbb{N}$, the bounded continuous functions over \mathbb{R} , and $\$X$, if X is Polish; the rationals \mathbb{Q} are **not** Polish. A measurable image of a Polish space is called an *analytic* space. Polish spaces, and to some extent, analytic spaces, have very convenient measure theoretic properties (for a general and accessible account, see [24], for a discussion tailored towards the purposes discussed here, see [9, Sections 4.3, 4.4]).

We note that a Kripke model over a Polish space is always small. This is so because the equivalence relation induced by the logic is *smooth*, i.e., countably generated, because the logic has only countably many formulas. The observation follows then from some general properties about smooth equivalence relations on Polish spaces [9, Proposition 4.4.26].

Now let (X, K) and (Y, L) be small models which are logically equivalent. We want to show that they are behaviorally equivalent, so we have to construct a mediator. Let α_K resp. α_L be the

Small
Kripke
model

Polish,
analytic

Smooth

corresponding equivalence relations with classes $[\cdot]_K$ and ρ_K as factor map, similarly for L . Define

$$\begin{aligned}\mathfrak{R} &:= \{\langle x, y \rangle \mid \text{Th}_K(x) = \text{Th}_L(y)\}, \\ \mathfrak{R}_0 &:= \{\langle [x]_K, [y]_L \rangle \mid \langle x, y \rangle \in \mathfrak{R}\}\end{aligned}$$

Since the models are logically equivalent, \mathfrak{R} is both left and right total. We will show now that \mathfrak{R}_0 is the graph of a bi-measurable map $f : X/K \rightarrow Y/L$. From the construction of \mathfrak{R}_0 it is clear that f is a bijection, so we have to cater for measurability. This is done through the principle of good sets in conjunction with the π - λ -Theorem. Consider

$$\mathcal{D} := \{B \subseteq Y/L \text{ measurable} \mid f^{-1}[B] \subseteq X/K \text{ is measurable}\}.$$

Then \mathcal{D} is closed under complementation and under countable disjoint unions. Let φ be a formula, then

$$f^{-1}[\rho_L[\llbracket \varphi \rrbracket_L]] = \rho_K[\llbracket \varphi \rrbracket_K].$$

Thus $\rho_L[\llbracket \varphi \rrbracket_L] \in \mathcal{D}$ for each formula φ . Since $\{\rho_L[\llbracket \varphi \rrbracket_L] \mid \varphi \text{ is a formula}\}$ is a \cap -stable generator of the factor σ -algebra due to (Y, L) being small, we conclude from the π - λ -Theorem that \mathcal{D} equals the σ -algebra of all measurable sets of Y/L . This shows that f is measurable, the argumentation is exactly the same for f^{-1} . Consequently, \mathfrak{R}_0 is the graph of an isomorphism.

Now look at this diagram:

$$\begin{array}{ccc} X & & Y \\ \rho_K \downarrow & & \downarrow \rho_L \\ X/\alpha_K & \xlongequal{\quad} & Y/\alpha_L \end{array}$$

Since ρ_K as well as ρ_L are morphisms $(X, K) \rightarrow (X/\alpha_K, K_\alpha)$ resp. $(Y, L) \rightarrow (Y/\alpha_L, L_\alpha)$, we have established this counterpart to the Hennessy-Milner Theorem:

Proposition 7.5 *Small logically equivalent Kripke models are behaviorally equivalent. \dashv*

What about bisimilarity? Rutten's paper [21] gives a calculus of bisimilarity for coalgebras. Unfortunately the really interesting properties assume that the functor under consideration preserves weak pullbacks. This is not the case for the Giry functor, as the following example demonstrates. It presents a situation in which no semi-pullback exists. A first example in this direction was suggested in [22, Theorem 12]. It is based on the extension of Lebesgue measure to a σ -algebra which does contain the Borel sets of $[0, 1]$ augmented by a non-measurable set, and it shows that one can construct Markov transition systems which do not have a semi-pullback. The example below simplifies this by showing that one does not have to consider transition systems, but that a look at the measures on which they are based suffices.

Example 7.6 A morphism $f : (X, \mathcal{A}, \mu) \rightarrow (Y, \mathcal{B}, \nu)$ of measure spaces is an \mathcal{A} - \mathcal{B} -measurable map $f : X \rightarrow Y$ such that $\nu = \$ (f) (\mu)$. Since each finite measure can be viewed as a transition kernel, this is a special case of morphisms for transition kernels. If \mathcal{B} is a sub- σ -algebra of \mathcal{A} with μ an extension to ν , then the identity is a morphisms $(X, \mathcal{A}, \mu) \rightarrow (X, \mathcal{B}, \nu)$.

Denote Lebesgue measure on $([0, 1], \mathcal{B}([0, 1]))$ by λ . Assuming the Axiom of Choice, we know that there exists $W \subseteq [0, 1]$ with $\lambda_*(W) = 0$ and $\lambda^*(W) = 1$. Here λ_* and λ^* denote the inner resp.

outer measure associated with Lebesgue measure. The technical condition says that whenever we have a measurable set $P \subseteq W$, then $\lambda(P) = 0$ must hold, and whenever we have a measurable set Q with $W \subseteq Q$, then $\lambda(Q) = 1$. These conditions make sure that W is not in the universal completion of $[0, 1]$, which means that Lebesgue measure cannot be extended uniquely to it in a canonic way.

But we find other, less canonic extensions, actually, plenty of them. Denote by

$$\mathcal{A}_W := \sigma(\mathcal{B}([0, 1]) \cup \{W\})$$

the smallest σ -algebra containing the Borel sets of $[0, 1]$ and W . We find for each $\alpha \in [0, 1]$ a measure μ_α on \mathcal{A}_W which extends λ such that $\mu_\alpha(W) = \alpha$ by [9, Exercise 4.6].

Hence by the remark just made, the identity yields a morphism

$$f_\alpha : ([0, 1], \mathcal{A}_W, \mu_\alpha) \rightarrow ([0, 1], \mathcal{B}([0, 1]), \lambda).$$

Now let $\alpha \neq \beta$, then

$$([0, 1], \mathcal{A}_W, \mu_\alpha) \xrightarrow{f_\alpha} ([0, 1], \mathcal{B}([0, 1]), \lambda) \xleftarrow{f_\beta} ([0, 1], \mathcal{A}_W, \mu_\beta)$$

is a co-span of morphisms. We claim that this co-span does not have a semi-pullback. In fact, assume that (P, \mathcal{P}, ρ) with morphisms π_α and π_β is a semi-pullback, then $f_\alpha \circ \pi_\alpha = f_\beta \circ \pi_\beta$, so that $\pi_\alpha = \pi_\beta$, and $\pi_\alpha^{-1}[W] = \pi_\beta^{-1}[W] \in \mathcal{P}$. But then

$$\alpha = \mu_\alpha(W) = \rho(\pi_\alpha^{-1}[W]) = \rho(\pi_\beta^{-1}[W]) = \mu_\beta(W) = \beta.$$

This contradicts the assumption that $\alpha \neq \beta$. \spadesuit

The question whether behaviorally equivalent Kripke models are bisimilar was open for quite some time, until Desharnais, Edalat and Panangaden showed in [3, 12] that behaviorally equivalent Kripke models based on analytic spaces are bisimilar with an analytic mediating system. This result was sharpened in [5, 4]: if the contributing models are based on Polish spaces, there exists a mediator based on a Polish space as well. Interestingly, the proof techniques are very different. While the analytic case is dealt with using conditional expectations, which are known to exist in analytic spaces, the Polish case is based on a selection argument, which — strange enough — does not generalize to analytic spaces.

The following proposition summarizes the discussion (see [3] for the analytic case, and [9, Proposition 4.10.20] for the Polish case):

Proposition 7.7 *Let (X_i, K_i) be Kripke models over analytic spaces X_1, X_2 , and assume that (X, K) is a stochastic relation, where X is a second countable metric space. Assume that we have a co-span of morphisms $m_i : K_i \rightarrow K, i = 1, 2$, then there exists a stochastic relation M and*

morphisms $m_i^+ : M \rightsquigarrow K_i, i = 1, 2$ rendering this diagram commutative.

$$\begin{array}{ccc} M & \xrightarrow{m_1^+} & K_2 \\ m_2^+ \downarrow & & \downarrow m_2 \\ K_1 & \xrightarrow{m_1} & K \end{array}$$

The stochastic relation M is defined over an analytic space. If X_1 and X_2 are Polish, M can be defined over a Polish space. \dashv

Proposition 7.7 is the crucial step in establishing [9, Proposition 4.10.22]:

Proposition 7.8 *Logically equivalent Kripke models over analytic spaces are bisimilar. The mediating model is analytic again. If the contributing models are Polish, the mediator is Polish as well. \dashv*

Modification

These results are formulated for Kripke models for the basic modal language with the grammar

1

$$\varphi ::= \top \mid \varphi_1 \wedge \varphi_2 \mid \Diamond_r \varphi$$

with $r \in [0, 1] \cap \mathbb{Q}$. They generalize easily to a modal logic in which the modalities are given through $\langle a \rangle_r$ for some action $a \in A$. Here we associate with each action a a stochastic relation $K_a : X \rightsquigarrow X$. A Kripke model is then given through $(X, (K_a)_{a \in A})$, and a morphism $f : (X, (K_a)_{a \in A}) \rightarrow (Y, (L_a)_{a \in A})$ is then a measurable map $f : X \rightarrow Y$ such that $L_a \circ f = \$ (f) \circ K_a$ holds for all $a \in A$. Another modification addresses the introduction of primitive formulas. The formulas for the general modal logic now look as follows:

Modification

2

$$\varphi ::= p \mid \top \mid \varphi_1 \wedge \varphi_2 \mid \langle a \rangle_r \varphi$$

with $p \in \Psi$ a primitive formula, $a \in A$ an action, and $r \in [0, 1] \cap \mathbb{Q}$. A Kripke model is now given through $(X, (V_p)_{p \in \Psi}, (K_a)_{a \in A})$, where $V_p \subseteq X$ is a measurable subset for each $p \in \Psi$, indicating the set of states in which a primitive formula holds; accordingly, we put $\llbracket p \rrbracket := V_p$. A morphism $f : (X, (V_p)_{p \in \Psi}, (K_a)_{a \in A}) \rightarrow (Y, (W_p)_{p \in \Psi}, (L_a)_{a \in A})$ is defined as above with the additional requirement that $f^{-1}[W_p] = V_p$ holds for each $p \in \Psi$.

The grammar above will be the one to use in the sequel as a sort of shell, where we fill in specific sets of actions.

8 Stochastic Effectivity Functions

We will now look into the interpretation of game logics; the reader interested in computational aspects is referred to [18]. In terms of modal logics, the actions are games, so the modalities are not flat, but rather structured according to the grammar through which games are specified. We will first make some general remarks.

Angel and Demon play against each other, taking turns. The two person game is modelled by this grammar

Two person
game

$$\tau ::= \gamma \mid \tau^d \mid \tau_1 \cup \tau_2 \mid \tau_1 \cap \tau_2 \mid \tau_1; \tau_2 \mid \tau^* \mid \tau^\times$$

with $\gamma \in \Gamma$, the set of atomic games. Games can be combined in different ways. If τ and τ' are games, $\tau; \tau'$ is the sequential composition of τ_1 and τ_2 , so that plays τ first, then τ' . In the game $\tau \cup \tau'$, Angel has the first move and decides whether τ or τ' is to be played, then the chosen game is played; $\tau \cup \tau'$ is called the *angelic choice* between τ_1 and τ_2 . Similarly, in $\tau \cap \tau'$ Demon has the first move and decides whether τ or τ' is to be played; accordingly, $\tau \cap \tau'$ is the *demonic choice* between the games. In the game τ^* , game τ is played repeatedly, until Angel decides to stop; it is not said in advance how many times the game is to be played, but it has to stop at some time; this is called *angelic iteration*. Dually, Demon decides to stop for the game τ^\times ; this is called *demonic iteration*. Finally, the rôles of Angel and Demon are interchanged in the game τ^d , so all decisions made by Demon are now being made by Angel, and vice versa.

When writing down games, we assume for simplicity that composition binds tighter than angelic or demonic choice. We make these assumptions [20, 18]:

- ① $(\tau^d)^d$ is identical to τ (recall that \cdot^d indicates Angel and Demon switching rôles).
- ② Demonic choice can be represented through angelic choice: The game $\tau_1 \cap \tau_2$ coincides with the game $(\tau_1^d \cup \tau_2^d)^d$.
- ③ Similarly, demonic iteration can be represented through its angelic counterpart: $(\tau^\times)^d$ is equal to $(\tau^d)^*$,
- ④ Composition is right distributive with respect to angelic choice: Making a decision to play τ_1 or τ_2 and then playing τ should be the same as deciding to play $\tau_1; \tau$ or $\tau_2; \tau$, thus $(\tau_1 \cup \tau_2); \tau$ equals $\tau_1; \tau \cup \tau_2; \tau$.

Note that left distributivity would mean that a choice between $\tau; \tau_1$ and $\tau; \tau_2$ is the same as playing first τ then $\tau_1 \cup \tau_2$; this is a somewhat restrictive assumption, since the choice of playing τ_1 or τ_2 may be a decision made by Angel only after τ is completed [25, p. 191]. Thus we do not assume this in general (it can be shown, however, that in Kripke generated models these choices are in fact equivalent [9, Proposition 4.9.40]).

- ⑤ We assume similarly that $\tau^*; \tau_0$ equals $\tau_0 \cup \tau^*; \tau; \tau_0$. Hence when playing $\tau^*; \tau_0$ Angel may decide to play τ not at all and to continue with τ_0 right away, or to play τ^* followed by $\tau; \tau_0$. Thus $\tau^*; \tau_0$ expands to $\tau_0 \cup \tau; \tau_0 \cup \tau; \tau; \tau_0 \cup \dots$.
- ⑥ $(\tau_1; \tau_2)^d$ is the same as $\tau_1^d; \tau_2^d$.
- ⑦ Angelic and demonic choice are commutative and associative, composition is associative.

For arriving at an interpretation, some historic remarks are helpful, and in order. Parikh [18], and later Pauly [19] propose interpreting game logic through a neighborhood model. Assign to each primitive game γ and each player (Angel: A; Demon: D) a neighborhood relation $N_g^{(i)} \subseteq S \times \mathcal{Z}^S$ ($i \in \{A, D\}$) with the understanding that $xN_\gamma^{(i)}S$ indicates player i having a strategy in

$N_g^{(i)}$

state x to force a state in $S \subseteq X$. Here X is the set of states over which the game is interpreted. The fact that $xN_\gamma^{(i)}S$ is sometimes described by saying that player i is effective for S (with game γ in state x). It is desirable that $xN_g^{(i)}S$ and $S \subseteq S'$ imply $xN_g^{(i)}S'$ for all states x . We assume that the game is *determined*, i.e., that exactly one of the players has a winning strategy. Thus $S \subseteq X$ is effective for player A in state x if and only if $X \setminus S$ is not effective for player D in that state. Consequently,

Determined
game

$$xN_\gamma^{(D)}S \Leftrightarrow \neg(xN_g^{(A)}X \setminus S),$$

which in turn implies that we only have to cater for Angel. We will omit the superscript from the neighborhood relation N_γ . Define the map $H_\gamma : X \rightarrow \mathbf{2}^{2^X}$ upon setting

$$H_\gamma(x) := \{S \subseteq X \mid xN_\gamma S\},$$

then $H_\gamma(x)$ is for all $x \in X$ an upper closed subset of $\mathbf{2}^X$ from which relation N_γ can be recovered. This function is called the *effectivity function* associated with relation N_γ . From N_γ another map $\tilde{N}_\gamma : \mathbf{2}^X \rightarrow \mathbf{2}^X$ is obtained upon setting

$$\tilde{N}_\gamma(A) := \{x \in X \mid xN_\gamma A\} = \{x \in X \mid A \in H_\gamma(x)\}.$$

Thus state x is an element of $\tilde{N}_\gamma(A)$ iff Angel has a strategy to force the outcome A when playing γ in x ; \tilde{N}_γ is actually a natural transformation. The operations on games can be taken care of through this family of maps, e.g., one sets recursively

$$\begin{aligned}\tilde{N}_{\tau_1 \cup \tau_2}(A) &:= \tilde{N}_{\tau_1}(A) \cup \tilde{N}_{\tau_2}(A), \\ \tilde{N}_{\tau_1; \tau_2}(A) &:= (\tilde{N}_{\tau_1} \circ \tilde{N}_{\tau_2})(A), \\ \tilde{N}_{\tau^*} &:= \bigcup_{n \geq 0} \tilde{N}_{\tau^n}(A).\end{aligned}$$

This refers only to Angel, Demon is accommodated through $A \mapsto S \setminus N_\gamma(X \setminus A)$ for primitive game γ , and by the rules ① to ⑦ from above. The maps \tilde{N}_τ serve in Parikh's original paper as a basis for defining the semantics of game logic. They are in one-to-one correspondence with effectivity functions, hence effectivity functions are the main actors.

For a probabilistic interpretation of game logic, it turns out to be convenient to also use effectivity functions as maps to upper closed subsets. **But subsets of what?**

Portfolio,
requirements

We observe these requirements for the *portfolio*, i.e., for the sets comprising the effectivity function:

1. The elements of the sets should be probability measures. This is so because we want to force a distribution over the states, rather than a state proper.
2. The portfolio should consist of measurable sets, so that we can measure them.
3. Stochastic relations should be a special case, hence it should be possible to integrate them swiftly.

Hence we require measurable sets of probabilities as possible outcomes, but this is not enough. We will also impose a condition on measurability on the interplay between distributions on states

and reals for measuring the probabilities of sets of states. This will lead to the definition of a stochastic effectivity function.

Denote for a measurable space X the \ast - σ -algebra on $\$(X)$ by $\mathcal{B}(\$(X))$, and put

$$\mathbb{F}(X) := \{V \subseteq \mathcal{B}(\$(X)) \mid V \text{ is upper closed}\}.$$

A measurable map $f : X \rightarrow Y$ induces a map $\mathbb{F}(f) : \mathbb{F}(X) \rightarrow \mathbb{F}(Y)$ upon setting

$\mathbb{F}(X), \mathbb{F}(f)$

$$\mathbb{F}(f)(V) := \{W \in \mathcal{B}(\$(Y)) \mid \$(f)^{-1}[W] \in V\}$$

for $V \in \mathbb{F}(X)$, then clearly $\mathbb{F}(f)(V) \in \mathbb{F}(Y)$.

Note that $\mathbb{F}(X)$ has not been equipped with a σ -algebra, so the usual notion of measurability between measurable spaces cannot be applied. In particular, \mathbb{F} is not an endofunctor on the category of measurable spaces. We will not discuss functorial aspects of \mathbb{F} here, referring the reader to [8] instead.

It would be most convenient if we could work in a monad — after all, the semantics pertaining to composition of games is modelled appropriately using a composition operator, as demonstrated through the definition of $\tilde{N}_{\tau_1; \tau_2}$ above. Markov transition systems are based on the Kleisli morphisms for the Giry monad, and the functor assigning each set upper closed subsets of the power set form a monad as well, see page 14. So one might want to capitalize on the composition of these monads. Alas, it is well known that the composition of two monads is not necessarily a monad, so this approach does not work, and one has to resort to ad-hoc methods simulating the properties of a monad (or of a Kleisli tripel).

Preparing for this, we require some properties pertaining to measurability, when dealing with the composition of distributions when discussing composite games. This will be provided in the following way. Let H be a measurable subset of $\$(X) \times [0, 1]$ indicating a quantitative assessment of subprobabilities (a typical example could be

$$\{\langle \mu, q \rangle \mid \mu \in \beta(A, q), 0 \leq q \leq 1\}$$

Quantitative
aspect

for some measurable $A \subseteq X$). Fix some real q and consider the set

$$H_q := \{\mu \mid \langle \mu, q \rangle \in H\}$$

of all measures evaluated through q . We ask for all states s such that this set is effective for s . They should come from a measurable subset of X . It turns out that this is not enough, we also require the real components being captured through a measurable set as well — after all, the real component will be used to be averaged, i.e., integrated, over later on, so it should behave decently. This idea is captured in the following definition.

Stochastic
effectivity
function

Definition 8.1 *Call a map $P : X \rightarrow \mathbb{F}(X)$ t -measurable iff $\{\langle s, q \rangle \mid H_q \in P(s)\} \subseteq X \times [0, 1]$ is measurable whenever $H \subseteq \$(X) \times [0, 1]$ is measurable. A stochastic effectivity function P on a measurable space X is a t -measurable map $P : X \rightarrow \mathbb{F}(X)$.*

Each stochastic relation gives rise to an effectivity function; this is indicated in the next example. The converse question, viz., under what conditions a stochastic effectivity function is generated by a stochastic relation, is more interesting, but a bit more cumbersome to answer. For the sake of completeness we indicate a characterization in the Appendix, see Section 8.1.

Example 8.2 Let $K : X \rightsquigarrow X$ be a stochastic relation, then

$$P_K(s) := \{A \subseteq \mathcal{P}(X) \mid K(s) \in A\}$$

is a stochastic effectivity function. \mathbb{M}

The next example is a little more sophisticated. It converts a finite transition system over a finite state space into an effectivity function.

Example 8.3 Let $X := \{1, \dots, n\}$ for some $n \in \mathbb{N}$, and take the power set as a σ -algebra. Then $\mathcal{P}(X)$ can be identified with the compact convex set

$$\Pi_n := \{\langle x_1, \dots, x_n \rangle \mid x_i \geq 0 \text{ for } 1 \leq i \leq n, \sum_{i=1}^n x_i = 1\}.$$

Geometrically, Π_n is the convex hull of the unit vectors e_i , $1 \leq i \leq n$; here $e_i(i) = 1$, and $e_i(j) = 0$ if $i \neq j$ is the i -th n -dimensional unit vector. The weak- $*$ - σ -algebra is the Borel- σ -algebra $\mathcal{B}(\Pi_n)$ for the Euclidean topology on Π_n .

Assume we have a transition system \rightarrow_X on X , hence a relation $\rightarrow_X \subseteq X \times X$. Put $\text{succ}(s) := \{s' \in X \mid s \rightarrow_X s'\}$ as the set of a successor states for state s . Define for $s \in X$ the set of weighted successors

$$\kappa(s) := \{\sum_{s' \in \text{succ}(s)} \alpha_{s'} \cdot e_{s'} \mid \mathbb{Q} \ni \alpha_{s'} \geq 0 \text{ for } s' \in \text{succ}(s), \sum_{s' \in \text{succ}(s)} \alpha_{s'} = 1\}$$

and the upper closed set

$$P(s) := \{A \in \mathcal{B}(\Pi_n) \mid \kappa(s) \subseteq A\}$$

A set A is in the portfolio for P in state s if A contains all rational distributions on the successor states. Here we restrict our attention to these rational distributions, which are positive convex combinations of the unit vectors with rational coefficients.

Then P can be shown to be a stochastic effectivity function on X [9, Example 4.1.14]. Actually, I don't know what happens when we admit real coefficients (things may become very complicated, then, since measurability might get lost). \mathbb{M}

8.1 Effectivity Functions vs. Stochastic Relations

The tools for investigating the converse to Example 8.2 come from the investigation of deduction systems for probabilistic logics. In fact, we are given a set of portfolios and want to know under which conditions this set is generated from a single subprobability. The situation is roughly similar to the one observed with deduction systems, where a set of formulas is given, and one wants to know whether this set can be constructed as valid under a suitable model. Because of the similarity, we may take some inspiration from the work on deduction systems, adapting the

approach proposed by R. Goldblatt [13]. Goldblatt works with formulas while we are interested foremost in families of sets; this permits a technically somewhat lighter approach in the present scenario.

Let S be a measurable space of states; we will deal with the measurable sets \mathcal{A} of S explicitly, so they are no longer swept under the carpet.

We first have a look at a relation $R \subseteq [0, 1] \times \mathcal{A}$ which models bounding probabilities from below. Intuitively, $\langle r, A \rangle \in R$ is intended to characterize the set $\beta(A, \geq r)$.

Definition 8.4 $R \subseteq [0, 1] \times \mathcal{A}$ is called a characteristic relation on S iff these conditions are satisfied

$$\begin{array}{ll}
 \textcircled{1} \frac{\langle r, A \rangle \in R, A \subseteq B}{\langle r, B \rangle \in R} & \textcircled{2} \frac{\langle r, A \rangle \in R, r \geq s}{\langle s, A \rangle \in R} \\
 \textcircled{3} \frac{\langle r, A \rangle \notin R, \langle s, B \rangle \notin R, r + s \leq 1}{\langle r + s, A \cup B \rangle \notin R} & \textcircled{4} \frac{\langle r, A \cup B \rangle \in R, \langle s, A \cup (S \setminus B) \rangle \in R, r + s \leq 1}{\langle r + s, A \rangle \in R} \\
 \textcircled{5} \frac{\langle r, A \rangle \in R, r + s > 1}{\langle s, S \setminus A \rangle \notin R} & \textcircled{6} \frac{\langle r, \emptyset \rangle \in R}{r = 0} \\
 \textcircled{7} \frac{A_1 \supseteq A_2 \supseteq \dots, \forall n \in \mathbb{N} : \langle r, A_n \rangle \in R}{\langle r, \bigcap_{n \geq 1} A_n \rangle \in R} & \textcircled{8} \langle 1, S \rangle \in R
 \end{array}$$

The conditions ① and ② make sure that bounding from below is monotone both in its numeric and in its set valued component. By ③ and ④ we cater for sub- and superadditivity of the characteristic relation, condition ⑥ sees to the fact that the probability for the impossible event cannot be bounded from below but through 0, and finally ⑦ makes sure that if the members of a decreasing sequence of sets are uniformly bounded below, then so is its intersection. These conditions are adapted from the S-axioms for T-deduction systems in [13, Section 4]. An exception is ⑦ which is weaker than the Countable Additivity Rule in [13, Definition 4.4]; we do not need a rule as strong as the latter one because we work with sets, hence we can deal with descending chains of sets directly.

We show that each characteristic relation defines a probability measure; the proof follows *mutatis mutandis* [13, Theorem 5.4].

Proposition 8.5 Let $R \subseteq [0, 1] \times \mathcal{A}$ be a characteristic relation on S , and define for $A \in \mathcal{A}$

$$\mu_R(A) := \sup\{r \in [0, 1] \mid \langle r, A \rangle \in R\}.$$

Then μ_R is a probability measure on \mathcal{A} . \dashv

Proof 1. ⑥ implies that $\mu_R(\emptyset) = 0$, and μ_R is monotone because of ①. It is also clear that $\mu_R(A) \leq 1$ always holds. We obtain from ② that $\langle s, A \rangle \notin R$, whenever $s \geq r$ with $\langle r, A \rangle \notin R$. Trivially, ⑧ implies that $\mu_R(S) = 1$.

2. Let $A_1, A_2 \in \mathcal{A}$ be arbitrary. Then

$$\mu_R(A_1 \cup A_2) \leq \mu_R(A_1) + \mu_R(A_2).$$

In fact, if $\mu_R(A_1) + \mu_R(A_2) < q_1 + q_2 \leq \mu_R(A_1 \cup A_2)$ with $\mu_R(A_i) < q_i$ ($i = 1, 2$), then $\langle q_i, A_i \rangle \notin R$

for $i = 1, 2$. Because $q_1 + q_2 \leq 1$, we obtain from ③ that $\langle q_1 + q_2, A_1 \cup A_2 \rangle \notin R$. By ② this yields $\mu_R(A_1 \cup A_2) < q_1 + q_2$, contradicting the assumption.

3. If A_1 and A_2 are disjoint, we observe first that $\mu_R(A_1) + \mu_R(A_2) \leq 1$. Assume otherwise that we can find $q_i \leq \mu_R(A_i)$ for $i = 1, 2$ with $q_1 + q_2 > 1$. Because $\langle q_1, A_1 \rangle \in R$ we conclude from ⑤ that $\langle q_2, S \setminus A_1 \rangle \notin R$, hence $\langle q_2, A_2 \rangle \notin R$ by ①, contradicting $q_2 \leq \mu_R(A_2)$.

This implies that

$$\mu_R(A_1) + \mu_R(A_2) \leq \mu_R(A_1 \cup A_2).$$

Assuming this to be false, we find $q_1 \leq \mu_R(A_1), q_2 \leq \mu_R(A_2)$ with

$$\mu_R(A_1 \cup A_2) < q_1 + q_2 \leq \mu_R(A_1) + \mu_R(A_2).$$

Because $\langle q_1, A_1 \rangle \in R$, we find $\langle q_1, (A_1 \cup A_2) \cap A_1 \rangle \in R$, because $\langle q_2, A_2 \rangle \in R$ we see that $\langle q_2, (A_1 \cup A_2) \cap (S \setminus A_1) \rangle \in R$ (note that $(A_1 \cup A_2) \cap A_1 = A_1$ and $(A_1 \cup A_2) \cap (S \setminus A_1) = A_2$, since $A_1 \cap A_2 = \emptyset$). From ④ we infer that $\langle q_1 + q_2, A_1 \cup A_2 \rangle \in R$, so that $q_1 + q_2 \leq \mu_R(A_1 \cup A_2)$, which is a contradiction.

Thus we have shown that μ_R is additive.

4. From ⑦ it is obvious that

$$\mu_R(A) = \inf_{n \in \mathbb{N}} \mu_R(A_n),$$

whenever $A = \bigcap_{n \in \mathbb{N}} A_n$ for the decreasing sequence $(A_n)_{n \in \mathbb{N}}$ in \mathcal{A} . \dashv

We relate $Q \in \mathbb{F}(S)$ to the characteristic relation R on S by comparing $\beta(A, \geq q) \in Q$ with $\langle q, A \rangle \in R$ by imposing a syntactic and a semantic condition. They will be shown to be equivalent.

Definition 8.6 $Q \in \mathbb{F}(S)$ is said to satisfy the characteristic relation R on S ($Q \vdash R$) iff we have

$$\langle q, A \rangle \in R \Leftrightarrow \beta(A, \geq q) \in Q$$

for any $q \in [0, 1]$ and any $A \in \mathcal{A}$.

This is a syntactic notion. Its semantic counterpart reads like this:

Definition 8.7 Q is said to implement $\mu \in \mathbb{S}(S)$ iff

$$\mu(A) \geq q \Leftrightarrow \beta(A, q) \in Q$$

for any $q \in [0, 1]$ and any $A \in \mathcal{A}$. We write this as $Q \models \mu$.

Note that $Q \models \mu$ and $Q \models \mu'$ implies

$$\forall A \in \mathcal{A} \forall q \geq 0 : \mu(A) \geq q \Leftrightarrow \mu'(A) \geq q.$$

Consequently, $\mu = \mu'$, so that the measure implemented by Q is uniquely determined.

We will show now that syntactic and semantic issues are equivalent: Q satisfies a characteristic relation if and only if it implements the corresponding measure. This will be used in a moment for a characterization of those game frames which are generated from Kripke frames.

Proposition 8.8 $Q \vdash R$ iff $Q \models \mu_R$.

Proof “ $Q \vdash R \Rightarrow Q \models \mu_R$ ”: Assume that $Q \vdash R$ holds. It is then immediate that $\mu_R(A) \geq r$ iff $\beta(A, \geq r) \in Q$.

“ $Q \models \mu_R \Rightarrow Q \vdash R$ ”: If $Q \models \mu_R$ for relation $R \subseteq [0, 1] \times \mathcal{A}$, we establish that the conditions given in Definition 8.4 are satisfied [9, Proposition 4.1.23].

1. Let $\beta(A, \geq r) \in Q$ and $A \subseteq B$, thus $\mu_R(A) \geq r$, hence $\mu_R(B) \geq r$, which in turn implies $\beta(B, \geq r) \in Q$. Hence ① holds. ② is established similarly.
2. If $\mu_R(A) < r$ and $\mu_R(B) < s$ with $r + s \leq 1$, then $\mu_R(A \cup B) = \mu_R(A) + \mu_R(B) - \mu_R(A \cap B) \leq \mu_R(A) + \mu_R(B) < r + s$, which implies ③.
3. If $\mu_R(A \cup B) \geq r$ and $\mu_R(A \cup (S \setminus B)) \geq s$, then $\mu_R(A) = \mu_R(A \cup B) + \mu_R(A \cup (S \setminus B)) - \mu_R(B) \geq r + s - \mu_R(B) \geq r$, hence ④.
4. Assume $\mu_R(A) \geq r$ and $r + s > 1$, then $\mu_R(S \setminus A) = \mu_R(S) - \mu_R(A) < p$, thus ⑤ holds.
5. If $\mu_R(\emptyset) \geq r$, then $r = 0$, yielding ⑥.
6. Finally, if $(A_n)_{n \in \mathbb{N}}$ is decreasing with $\mu_R(A_n) \geq r$ for each $n \in \mathbb{N}$, then it is plain that $\mu_R(\bigcap_{n \in \mathbb{N}} A_n) \geq r$. This implies ⑦.

⊢

This permits a complete characterization of those stochastic effectivity functions which are generated through stochastic relations.

Proposition 8.9 *Let P be a stochastic effectivity frame on state space S . Then these conditions are equivalent*

1. *There exists a stochastic relation $K : S \rightsquigarrow S$ such that $P = P_K$.*
2. *$R(s) := \{\langle r, A \rangle \mid \beta(A, \geq r) \in P(s)\}$ defines a characteristic relation on S with $P(s) \vdash R(s)$ for each state $s \in S$.*

Proof “1 \Rightarrow 2”: Fix $s \in S$. Because $\beta(A, \geq r) \in P_K(s)$ iff $K(s)(A) \geq r$, we see that $P(s) \models K(s)$, hence by Proposition 8.8 $P(s) \vdash R(s)$.

“2 \Rightarrow 1”: Define $K(s) := \mu_{R(s)}$, for $s \in S$, then $K(s)$ is a subprobability measure on \mathcal{A} . We show that $K : S \rightsquigarrow S$. Let $G \subseteq \mathcal{S}(S)$ be a $*$ -measurable set, then $G \times [0, 1] \subseteq \mathcal{S}(S) \times [0, 1]$ is measurable, hence the measurability condition on P yields that

$$K^{-1}[G] = \{s \in S \mid K(s) \in G\} = \{s \in S \mid G \in P(s)\}$$

is a measurable subset of S , because

$$\{\langle s, q \rangle \mid (G \times [0, 1])_q \in P(s)\} = \{s \in S \mid G \in P(s)\} \times [0, 1] \subseteq S \times [0, 1] \text{ is measurable.}$$

⊢

8.2 Game Frames

We define *game frames* similar to Kripke frames as being comprised of a state space and the maps which indicate the actions to be taken.

Definition 8.10 A game frame $\mathcal{G} = (S, (P_\gamma)_{\gamma \in \Gamma})$ has a measurable space S of states and a t -measurable map $P_\gamma : S \rightarrow \mathcal{F}(S)$ for each primitive game $\gamma \in \Gamma$.

Game frame

Now that we have game frames, we can do some interesting things, e.g., use them for the interpretation of game logics (well, nearly). What we do first is to define recursively a set valued function $\Omega_{\mathcal{G}}(\tau \mid A, q)$ with the intention to describe the set of states for which Angel upon playing game τ has a strategy of reaching a state in set A with probability greater than q . Assume that $A \in \mathcal{B}(S)$ is a measurable subset of S , and $0 \leq q < 1$, and define for $0 \leq k \leq \infty$

$\Omega_{\mathcal{G}}(\tau \mid A, q)$

$$Q^{(k)}(q) := \{\langle a_1, \dots, a_k \rangle \in \mathbb{Q}^k \mid a_i \geq 0 \text{ and } \sum_{i=1}^k a_i \leq q\}.$$

as the set of all non-negative rational k -tuples resp. sequences the sum of which does not exceed q .

- ❶ Let $\gamma \in \Gamma$ be a primitive game, then put

$\gamma \in \Gamma$

$$\Omega_{\mathcal{G}}(\gamma \mid A, q) := \{s \in S \mid \beta(A, q) \in P_\gamma(s)\},$$

in particular

$$\Omega_{\mathcal{G}}(\epsilon \mid A, q) = \{s \in S \mid \delta_s(A) \geq q\} = A.$$

Thus $s \in \Omega_{\mathcal{G}}(\gamma \mid A, q)$ iff Angel has $\beta(A, q)$ in its portfolio when playing γ in state s . This implies that the set of all state distributions which evaluate at A with a probability greater than q can be effected by Angel in this situation. If Angel does not play at all, hence if the game γ equals ϵ , nothing is about to change, which means

$$\Omega_{\mathcal{G}}(\epsilon \mid A, q) = \{s \mid \delta_s \in \beta(A, q)\} = A,$$

as expected.

- ❷ Let τ be a game, then

τ^d

$$\Omega_{\mathcal{G}}(\tau^d \mid A, q) := S \setminus \Omega_{\mathcal{G}}(\tau \mid S \setminus A, q).$$

The game is determined, thus Demon can reach a set of states iff Angel does not have a strategy for reaching the complement. Consequently, upon playing τ in state s , Demon can reach a state in A with probability greater than q iff Angel cannot reach a state in $S \setminus A$ with probability greater q .

Illustrating, let us assume for the moment that $P_\gamma = P_{K_\gamma}$, i.e., the special case in which the effectivity function for the primitive game $\gamma \in \Gamma$ is generated from a stochastic relation K_γ .

Then

$$s \in \Omega_{\mathcal{G}}(\gamma^d \mid A, q) \Leftrightarrow s \notin \Omega_{\mathcal{G}}(\gamma \mid S \setminus A, q) \Leftrightarrow K_{\gamma}(s)(S \setminus A) \leq q.$$

In general,

$$s \in \Omega_{\mathcal{G}}(\gamma^d \mid A, q) \Leftrightarrow \beta(S \setminus A, q) \notin P_{\gamma}(s)$$

for $\gamma \in \Gamma$. This is exactly what one would expect in a determined game.

- ③ Assume s is a state such that Angel has a strategy for reaching a state in A when playing the game $\tau_1 \cup \tau_2$ with probability not greater than q . Then Angel should have a strategy in s for reaching a state in A when playing game τ_1 with probability not greater than a_1 and playing game τ_2 with probability not greater than a_2 such that $a_1 + a_2 \leq q$. Thus

$\tau_1 \cup \tau_2$

$$\Omega_{\mathcal{G}}(\tau_1 \cup \tau_2 \mid A, q) := \bigcap_{a \in Q^{(2)}(q)} (\Omega_{\mathcal{G}}(\tau_1 \mid A, a_1) \cup \Omega_{\mathcal{G}}(\tau_2 \mid A, a_2)).$$

- ④ Right distributivity of composition over angelic choice translates to this equation.

$(\tau_1 \cup \tau_2); \tau$

$$\Omega_{\mathcal{G}}((\tau_1 \cup \tau_2); \tau \mid A, q) := \Omega_{\mathcal{G}}(\tau_1; \tau \cup \tau_2; \tau \mid A, q).$$

- ⑤ If $\gamma \in \Gamma$, put

$\gamma; \tau$

$$\Omega_{\mathcal{G}}(\gamma; \tau \mid A, q) := \{s \in S \mid G_{\tau}(A, q) \in P_{\gamma}(s)\},$$

where

$$G_{\tau}(A, q) := \{\mu \in \mathcal{S}(S) \mid \int_0^1 \mu(\Omega_{\mathcal{G}}(\tau \mid A, r)) \, dr > q\}.$$

Suppose that $\Omega_{\mathcal{G}}(\tau \mid A, r)$ is already defined for each $r \in [0, 1]$ as the set of states for which Angel has a strategy to effect a state in A through playing τ with probability greater than r . Given a distribution μ over the states, the integral $\int_0^1 \mu(\Omega_{\mathcal{G}}(\tau \mid A, r)) \, dr$ is the expected value for entering a state in A through playing τ for μ . The set $G_{\tau}(A, q)$ collects all distributions the expected value of which is greater than q . We ask for all states such that Angel has this set in its portfolio when playing γ in this state. Being able to select this set from the portfolio means that when playing γ and subsequently τ a state in A may be reached with probability greater than q .

- ⑥ The transformation for $\tau^*; \tau_0$ is obtained through a fairly direct translation of assumption ⑤. by repeated application of the rule ③. for angelic choice:

$\tau^*; \tau_0$

$$\Omega_{\mathcal{G}}(\tau^*; \tau_0 \mid A, q) := \bigcap_{a \in Q^{(\infty)}(q)} \bigcup_{n \geq 0} \Omega_{\mathcal{G}}(\tau^n; \tau_0 \mid A, a_{n+1})$$

with $\tau^n := \tau; \dots; \tau$ (n times).

We obtain for state spaces that are closed under the Souslin operations (see Appendix B).

Proposition 8.11 *Assume that the state space is closed under the Souslin operation, then $\Omega_{\mathcal{G}}(\tau \mid A, q)$ is a measurable subset of S for all $A \subseteq S$ measurable, for all games τ , and $0 \leq q \leq 1$. \dashv*

Note that universally complete measurable spaces or analytic sets, which are popular in some circles, are closed under this mysterious operation (in fact, you can represent each analytic set in a Polish space through a Souslin scheme of closed sets [9, Proposition 4.5.6]), but other important spaces like Polish spaces are not.

Finally, we introduce *game models*; they are what you expect. Take a game frame and add for each primitive formula the set of all states in which it is assumed to be valid.

Definition 8.12 *A game model $\mathcal{G} = (S, (P_\gamma)_{\gamma \in \Gamma}, (V_p)_{p \in \Psi})$ over measurable space S is given by a game frame $(S, (P_\gamma)_{\gamma \in \Gamma})$, and by a family $(V_p)_{p \in \Psi}$ of sets which assigns to each atomic statement a measurable set of state space S . We denote the underlying game frame by \mathcal{G} as well.*

Game
model

Define the validity sets for each formula recursively as follows:

$$\begin{aligned} \llbracket \top \rrbracket_{\mathcal{G}} &:= S \\ \llbracket p \rrbracket_{\mathcal{G}} &:= V_p, \text{ if } p \in \Psi \\ \llbracket \varphi_1 \wedge \varphi_2 \rrbracket_{\mathcal{G}} &:= \llbracket \varphi_1 \rrbracket_{\mathcal{G}} \cap \llbracket \varphi_2 \rrbracket_{\mathcal{G}} \\ \llbracket \langle \tau \rangle_q \varphi \rrbracket_{\mathcal{G}} &:= \Omega_{\mathcal{G}}(\llbracket \varphi \rrbracket_{\mathcal{G}} \mid \tau, q) \end{aligned}$$

Accordingly, we say that formula φ holds in state s ($\mathcal{G}, s \models \varphi$) iff $s \in \llbracket \varphi \rrbracket_{\mathcal{G}}$.

The definition of $\llbracket \langle \tau \rangle_q \varphi \rrbracket_{\mathcal{G}}$ has a coalgebraic flavor. Coalgebraic logics define the validity of modal formulas through special natural transformations (called *predicate liftings*) associated with the modalities [17, 23, 6]. You may wish to look up the brief discussion in [9, Section 2.7.3].

Proposition 8.13 *If state space S is closed under the Souslin operation, $\llbracket \varphi \rrbracket_{\mathcal{G}}$ is a measurable subset for all formulas φ . Moreover, $\{\langle s, r \rangle \mid s \in \llbracket \langle \tau \rangle_r \varphi \rrbracket_{\mathcal{G}}\}$ is a measurable subset of $S \times [0, 1]$.*

This shows that the transformations we consider do not leave the realm of measurability, provided the space is decent enough.

Proof The proof proceeds by induction on the formula φ . If $\varphi = p \in \Psi$ is an atomic proposition, then the assertion follows from $V_p \in \mathcal{B}(S)$. The induction step uses Proposition 8.11. \dashv

9 What To Do Next

The next question to discuss would be the expressivity of game models. This is technically somewhat involved, because the base mechanism is not exactly light footed. I refer you to [10]. There also the relationship to Kripke models is investigated and completely characterized. The bridge to stochastic nondeterminism through what is called *hit measurability* from [2] is constructed in [11] with some new results on bisimilarity.

A The π - λ -Theorem

The measure theoretic reason why we insist on our modal logics being closed under conjunctions is Dynkin's famous π - λ -Theorem:

Theorem A.1 *Let \mathcal{P} be a family of subsets of a set S which is closed under finite intersections. Then $\sigma(\mathcal{P})$ is the smallest class containing \mathcal{P} which is closed under complements and countable disjoint unions. \dashv*

For a proof, see [9, Theorem 1.6.30]. There you find also a discussion on its use, e.g., when establishing the equality of measures from generators of σ -algebras to the σ -algebra proper. The application to modal logics is immediate, given that conjunction of formulas translates to the intersection of the validity sets.

B The Souslin Operation

V^* denotes for a set V the set of all finite words with letters from V including the empty string ϵ . Let $\{A_s \mid s \in \mathbb{N}^*\}$ be a collection of subsets of a set X indexed by all finite sequences of natural numbers (a *Souslin scheme*), then the *Souslin operation* \mathfrak{A} on this collection is defined as

$$\mathfrak{A}(\{A_s \mid s \in \mathbb{N}^*\}) := \bigcup_{\alpha \in \mathbb{N}^{\mathbb{N}}} \bigcap_{n \in \mathbb{N}} A_{\alpha|n},$$

where $\alpha|n \in \mathbb{N}^*$ is just the word composed from the first n letters of the sequence α . This operation is intimately connected with the theory of analytic sets [16, 1, 24]. We obtain from [1, Proposition 1.10.5]:

Proposition B.1 *If S is a universally complete measurable or an analytic space, then its measurable sets are closed under the operation \mathfrak{A} . \dashv*

A discussion of the Souslin operation together with some technical tools associated with it is given in [9, Section 4.5].

Acknowledgements. The tutorial on which these notes are based was presented in December 2015 and January 2016 at the Department of Mathematics and Computer Science of the University of Udine. I want to thank Prof. Marino Miculan and Prof. Alberto Policritti for inviting me to present this tutorial to their students; Prof. Eugenio Omodeo provided as always some helpful comments, for which I want to thank him.

References

- [1] V. I. Bogachev. *Measure Theory*. Springer-Verlag, 2007.
- [2] P. R. D'Argenio, P. Sánchez Terraf, and N. Wolovick. Bisimulations for non-deterministic labelled Markov processes. *Math. Struct. Comp. Science*, 22(1):43–68, February 2012.

- [3] J. Desharnais, A. Edalat, and P. Panangaden. Bisimulation of labelled Markov processes. *Information and Computation*, 179(2):163 – 193, 2002.
- [4] E.-E. Doberkat. Semi-pullbacks and bisimulations in categories of stochastic relations. In *Proc. ICALP'03*, volume 2719 of *Lect. Notes Comp. Sci.*, pages 996 – 1007, Berlin, 2003. Springer-Verlag.
- [5] E.-E. Doberkat. Semi-pullbacks for stochastic relations over analytic spaces. *Math. Struct. Comp. Sci.*, 15:647 – 670, 2005.
- [6] E.-E. Doberkat. *Stochastic Coalgebraic Logic*. EATCS Monographs in Theoretical Computer Science. Springer-Verlag, Berlin, 2009.
- [7] E.-E. Doberkat. A note on the coalgebraic interpretation of game logic. *Rendiconti Ist. di Mat. Univ. di Trieste*, 42:191 – 204, 2010.
- [8] E.-E. Doberkat. Algebraic properties of stochastic effectivity functions. *J. Logic and Algebraic Progr.*, 83:339 – 358, 2014.
- [9] E.-E. Doberkat. *Special Topics in Mathematics for Computer Science: Sets, Categories, Topologies, Measures*. Springer International Publishing Switzerland, Cham, Heidelberg, New York, Dordrecht, London, December 2015.
- [10] E.-E. Doberkat. A stochastic interpretation of game logic. arXiv: 1403.7765, February 2016.
- [11] E.-E. Doberkat and P. Sànchez Terraf. Stochastic nondeterminism and effectivity functions. *J. Logic and Computation*, doi: 10.1093/logcom/exv049 (arxiv: 1405.7141), 2015.
- [12] A. Edalat. Semi-pullbacks and bisimulations in categories of Markov processes. *Math. Struct. Comp. Science*, 9(5):523 – 543, 1999.
- [13] R. Goldblatt. Deduction systems for coalgebras over measurable spaces. *Journal of Logic and Computation*, 20(5):1069 – 1100, 2010.
- [14] M. Hennessy and R. Milner. On observing nondeterminism and concurrency. In *Proc. ICALP'80*, number 85 in *Lect. Notes Comp. Sci.*, pages 395 – 409, Berlin, 1980. Springer-Verlag.
- [15] B. Jacobs. *Introduction to coalgebra: Towards Mathematics of states and observations*. Cambridge University Press, Cambridge, UK, (to be published) edition, 2016.
- [16] T. Jech. *Set Theory*. Springer-Verlag (The Third Millennium Edition), Berlin, Heidelberg, New York, 2006.
- [17] L. M. Moss. Coalgebraic logic. *Ann. Pure Appl. Logic*, 96:277 – 317, 1999.
- [18] R. Parikh. The logic of games and its applications. In M. Karpinski and J. van Leeuwen, editors, *Topics in the Theory of Computation*, volume 24, pages 111–140. Elsevier, 1985.
- [19] M. Pauly. Game logic for game theorists. Technical Report INS-R0017, CWI, Amsterdam, 2000.

- [20] M. Pauly and R. Parikh. Game logic — an overview. *Studia Logica*, 75:165 – 182, 2003.
- [21] J. J. M. M. Rutten. Universal coalgebra: a theory of systems. *Theor. Comp. Sci.*, 249(1):3 – 80, 2000. Special issue on modern algebra and its applications.
- [22] P. Sánchez Terraf. Unprovability of the logical characterization of bisimulation. *Information and Computation*, 209(7):1048 – 1056, 2011.
- [23] L. Schröder. Expressivity of coalgebraic modal logic: the limits and beyond. *Theor. Comp. Sci.*, 390:230–247, 2008.
- [24] S. M. Srivastava. *A Course on Borel Sets*. Graduate Texts in Mathematics. Springer-Verlag, Berlin, 1998.
- [25] J. van Benthem. Logic games are complete for game logics. *Studia Logica*, 75:183 – 203, 2003.
- [26] Y. Venema. Algebras and co-algebras. In P. Blackburn, J. van Benthem, and F. Wolter, editors, *Handbook of Modal Logic*, volume 3 of *Studies in Logic and Practical Reasoning*, pages 331–426. Elsevier, Amsterdam, 2007.